REAL-TIME ANALYTICS FOR RESOURCE MOBILIZATION IN ENGINEERING DIVERSITY HASHTAG CAMPAIGNS

by

Habib Karbasian
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
In Partial fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Information Technology

Committee:

____________________	Dr. Aditya Johri, Dissertation Director
____________________	Dr. Hemant Purohit, Co Advisor
____________________	Dr. Huzefa Rangwala, Committee Member
____________________	Dr. Liang Zhao, Committee Member
____________________	Dr. Deborah Goodings, Associate Dean
____________________	Dr. Kenneth S. Ball, Dean, The Volgenau School of Engineering

Date: ____________________ Spring 2020
George Mason University
Fairfax, VA

Real-time Analytics for Resource Mobilization in
Engineering Diversity Hashtag Campaigns

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Habib Karbasian
Master of Science
University of Tehran, 2008
Bachelor of Science
University of Tehran, 2005

Director: Dr. Aditya Johri, Professor
Department of Information Sciences and Technology

Spring 2020
George Mason University
Fairfax, VA

Dedication

I dedicate this dissertation to my beloved family, Mehdi Karbasian, Shohreh Karimnejad and Hamed Karbasian whose constantly emotional and financial support have always eased my rough path toward PhD. Undoubtedly, I would have not been here if my father hadn't put me in right directions and my mom hadn't filled me with indefinite love.

Acknowledgments

I would like to thank my advisor, Dr. Aditya Johri, and co-advisor, Dr. Hemant Purohit, who made this work possible by positive and continuous support with insightful feedback to strengthen my dissertation. During my research journey, I have learned a lot about how to conduct an independent research from scratch to final phase. I will definitely apply my skills throughout my life and I will try my best to help my peers to improve.

I also would like to extend my gratitude toward my committe members: Dr. Huzefa Rangwala and Dr. Liang Zhao for their intricit and in-depth comments to help me have a wholesome dissertation.

I am also thrilled to share this accomplishment with my close friends who accompanied me in fun and rough times during my PhD specially the last phase of dissertation where support group cannot be replaced with anything else. I like to start my appreciation with Ali Mirzaeian who reviewed and edited parts of my dissertation and presentation, which without his help, the quality of my work would have not been the same. Also I used a lot of help from a close friend of mine, Abbas Zaghari, in Fairfax who has always made feel special and empowered by his positive viewpoints. My long time friend, Siavash Kayal, with whom I had long phone conversations about our shared past experiences from University of Tehran reminded me of where I started my dream and how I am making it happen. Last but not least, Mohammad Ansari, my close friend living in Iran, played a huge and important role in my success to make progress. He always said, *"There is no other option in life but winning!"*.

"The work presented here was supported in part by U.S. National Science Foundation Awards#: DUE-**1707837** and DUE-**1712129**. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the funding agencies."

Table of Contents

List of Tables

List of Figures

Abstract

REAL-TIME ANALYTICS FOR RESOURCE MOBILIZATION IN
ENGINEERING DIVERSITY HASHTAG CAMPAIGNS

Habib Karbasian, PhD

George Mason University, 2020

Dissertation Director: Dr. Aditya Johri

To counter the lack of diversity within the engineering profession and to promote equity, many activists utilize social media platforms to raise awareness of the issue. One of the goals of these activism campaigns is to break down stereotypes about who is an engineer and to make the diverse participation visible within the profession. Although prior work indicates that social media hashtag campaigns help promote diversity, it also outlines several challenges that need to be addressed for efficiency and increased inclusiveness. In particular, more work is needed to understand the initial phases of these campaigns because during that stage, efficient resource mobilization is needed to attract participants and make the campaign a success. For resource mobilization, campaigns need to identify and attract influential users and create a message that resonates with participants in real-time. By sharing effective messages, influential users can thus create momentum for the campaign. This dissertation research develops two analytical frameworks to assist in

finding out (a) which types of users can be more influential for educating and informing people in real-time, and (b) what kind of messages are deemed more attractive and inspirational to the targeted audience in real-time. To develop the frameworks and to test them, data from two Twitter hashtag campaigns is used: #ILookLikeAnEngineer (19492 tweets over the course of two months) and #WomenInEngineering (8225 tweets over the course of 16 months).

The first part of this dissertation addresses the problem of identifying user types in these hashtag campaigns in real-time. Our preliminary work showed that organizations and individuals both played a significant role in promoting these campaigns goals but in different ways, both quantitatively and qualitatively. Therefore, it was important to distinguish different kinds of users with this classification in mind. This problem was solved by advancing an analytical framework to classify users into individuals (male or female) or organizations in real-time. This framework uses the multimodality of the information available in one tweet from a given user, such as user's profile picture, user's name, user's network metadata, linguistic and psychological characteristics of the tweet, and determines the user type of the given tweet. The proposed framework with real-time features outperformed other baselines with more than 6.65% accuracy increase for ILookLikeAnEngineer and showed that organizations were one of the influential users after female participants.

The second part of this dissertation addresses the problem of identifying messages that resonate with users. Based on prior work in the literature, it was hypothesized that some latent features, i.e.: the cluster of various co-occurring hashtags along with relevant topics and relatable sentiments in the messages can be deciding factors to attract like-minded users. An analytical framework was developed that utilizes these features to predict if a given tweet will be retweeted in real-time. The proposed framework works with real-time data and outperformed other baselines with 6.61% and 8.25% accuracy increase respectively for ILookLikeAnEngineer and WomenInEngineering.

Chapter 1: Introduction

Figure 1.1 illustrates the big picture of this dissertation. It first provides an overview of STEM and its regarding issues, one of which is the focus of this work, engineering diversity in education and workforce [1] . Then it recommends a solution to tackle the issue by creating a positive and empowering culture for underrepresented minorities in form of raising awareness, promoting role models and breaking stereotypes in engineering fields. But the main issue is the lack of efficient and real-time messaging mechanism to provide right information at the right time for the targeted audience.

Social movement [1] is one of the avenues to bring social changes to the existing issues. Using social media platforms as a main resource for social movements, online activism with the help of hashtag, hashtag activism campaign [1] , provides a channel to spread the information to the targeted audience and it is able to address the messaging issue in engineering diversity efficiently.

Motivated by *ILookLikeAnEngineer* campaign [1] as a case study of engineering diversity solution, this dissertation later will review the important challenges for hashtag activism campaigns (*resource mobilization, activities and leadership*) and propose solutions afterwards. Next it will form the overarching question to address all of the challenges which later will be broken down to two research questions to find out 1) who can be more influential with educating and informing people in real-time, 2) what kind of messages are

[1] It will be further studied in "background and related works" section.

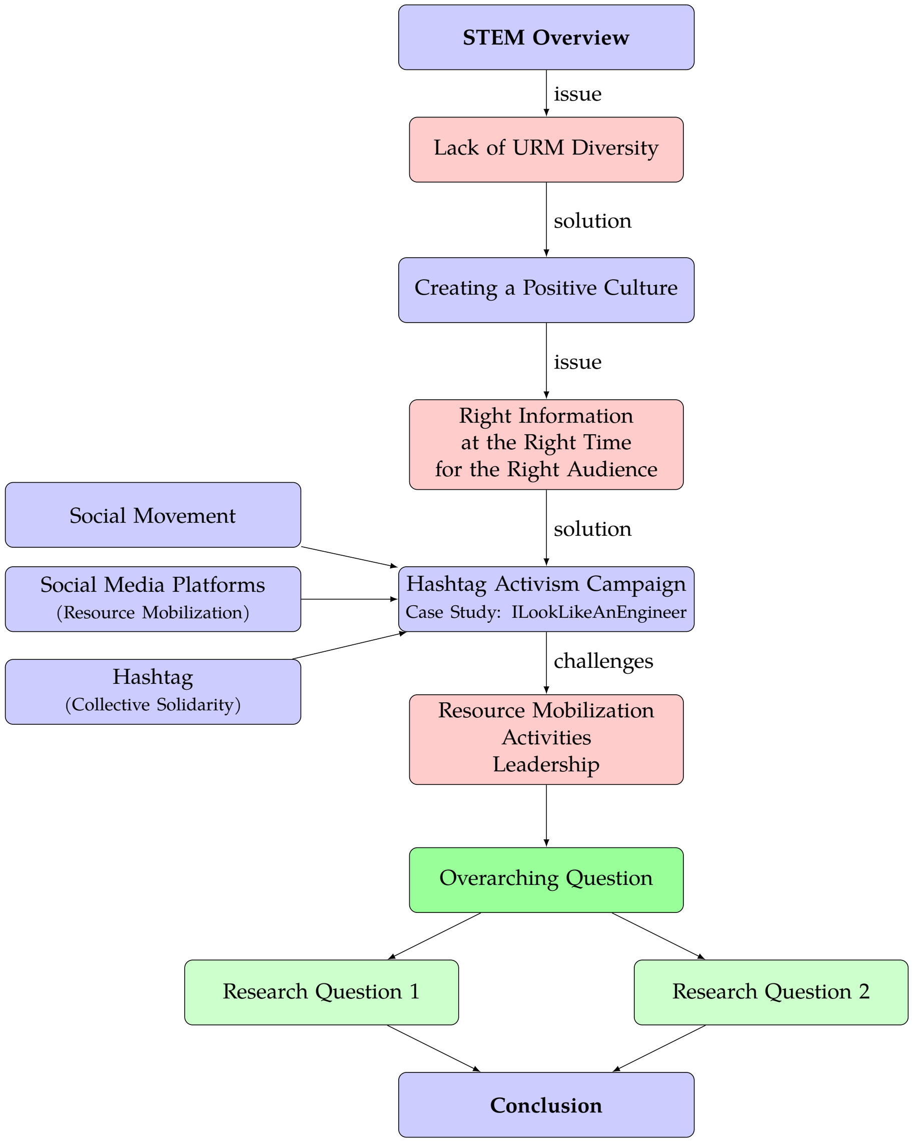

Figure 1.1: Summary Figure of the Dissertation

deemed more attractive and inspirational to the targeted audience in real-time. Finally conclusion will summarize the findings, future works and applications.

1.1 STEM Overview and Issues

Science, Technology, Engineering, and Mathematics (STEM) has become one of the focal points in educational activities in US across all grade levels (from pre-school to post-doctorate) and to both formal (e.g., classrooms) and informal (e.g., after school programs) experiences. To maintain its competitive edge across the world, US has prioritized STEM education in its policy making specially funding appropriation which goes beyond $4 billion annually [Malik et al., 2018]. According to the Congressional Research Service STEM primer [Gonzalez and Kuenzi, 2012], there are between 105 and 252 STEM education programs or activities at 13 to 15 federal agencies; the key agencies involved in the effort are Department of Education, National Science Foundation, and Health and Human Services. Given that the interests are broad and federal efforts are spread across multiple agencies, there is a concern with perceived duplication of effort and a lack of coordination in the federal effort. Therefore, efforts to improve the coordination have gained prominence in recent years.

The data currently available about STEM education paints a "complicated" picture [Gonzalez and Kuenzi, 2012]. According to many indicators, overall graduate enrollments in science and engineering (S&E) grew 35% over the last decade; S&E enrollments for Hispanic/Latino, American Indian/Alaska Native, and African American students (all of whom are generally underrepresented [2] in S&E) grew by 65%, 55%, and 50%, respectively. Yet, concerns remain about persistent academic achievement gaps between various demographic groups.

[2] Underrepresented minorities (URM) can be categorized as racial or gender-based. The focus on this research is to look further in gender disparity in STEM.

On the other hand, engineering has been also one of the most male-dominated professions in the world. While other professions such as law and medicine have now achieved gender parity at student level, at least in many industrially advanced countries, the number of women entering the engineering profession remains very low. In fact, the percentage of women starting to study in engineering has actually decreased over the last few years so that it is now less than 20 percent in several western countries [ASE, 2009,UKR, 2010]. These percentages are even lower for women working as professional engineers, and are frequently closer to 10% or less [Nat, 2007]. This low retention and acceptance in engineering workforce have been attributed to the characteristics of a masculine culture [Bastalich et al., 2003,McIlwee and Robinson, 1992,Miller, 2004,Powell et al., 2009]. Although women are as competent as their male colleagues at the technical dimensions of engineering (and, in fact, often do very well in their studies and obtain graduate positions quite readily), the expectations and processes in engineering organizations constitute a real problem for women's careers [Evetts, 1998,McIlwee and Robinson, 1992,Miller, 2004].

To address the lack of URM diversity in engineering fields, policies and campaigns were introduced and implemented to achieve equity for women for more than two decades. These policies have included affirmative action, equal employment opportunity, managing diversity and gender mainstreaming. Yet the gains have not been substantial. These policies help the system to attract and retain the underrepresented more likely than before but the issue is how to create a positive and inspiring culture targeted toward these groups to make them feel empowered and keep them encouraged in their educational and professional development. Malcom et al. (2013) [Malcom-Piqueux and Malcom, 2013] suggest that greater awareness of careers, role models, after- and out-of-school experiences are critical to address this issue. They argue that *"students and their families need encouragement* *and access to information at a much earlier stage than has typically been provided, through exposure to role models who look like them, information about the kinds of jobs done by persons with preparation in engineering, and examples to dispel the idea that engineering is solitary work (pg. 32)* [*Malcom-Piqueux and Malcom*, *2013*]." They further suggest that efforts to increase the participation of URM in engineering education and careers start outside of school, through programs such as Expanding Your Horizons for female students and Mathematics, Engineering, Science Achievement (MESA), Sally Ride Science and AAAS' Spark Club [3], After School Alliance, and others but few parents are aware of these programs. All the issues outlined here are easy to address if we are able to provide them with the right information at the right time. Furthermore, the efforts we make will be even more fruitful if we are able to create communication channels that facilitates the messaging mechanism, through which we can easily access the right target population at the right time.

After the introduction of Web 2, interactive web, social media platforms have become the center of public discourse and communication of all kind. These platforms, such as Facebook, Instagram, and Twitter, can help us understand the events and activities that are of concern and of issues that matter to us. Due to the increase usage and access to mobile systems hence using these apps on them, people continue to use these platforms to interact with each other leading to a constantly growing databases that are extremely valuable to mine and understand. Almost 87% of the American population now participates in some form of social media activity. Social media provide powerful indicators of social action as they allow individual and small groups to find like-minded individuals and organizations and topics around which they can engage in. Social media data has

[3]Part of a suite of programs under NSF's Innovative Technology Experiences for Students and Teachers (ITEST)

been especially useful for researchers studying issues around health communication [Kotsenas et al., 2018], politics [Highfield, 2017], humanitarian crisis [Reuter and Kaufhold, 2018], social movements such as Arab Spring and Black Lives Matters [Rickford, 2016], [Hänska Ahy, 2016] and activism [Johri et al., 2018b]. These efforts suggest that another approach to bring the change to engineering diversity issue is to use online social media platforms to educate other people or share their anecdotal experiences. It can be achieved by individually initiated social movements addressing aforementioned issues like *#ILookLikeAnEngineer*, *#WomenInSTEM* and *#WomenInEngineering* on Twitter.

Studying and analyzing these engineering diversity case studies in hashtag activism campaigns shed light on the current status of the issue at hand and sometimes lessons and best practices are extracted that can be applied to other campaigns. But if we only focus on the case study analysis, we will never truly understand online social movement on general because the use and relevance of digital tools and tactics are constantly changing. The goal of this work is to move beyond and focus on underlying mechanics. While we continue to be inspired and fascinated by these examples for engineering diversity around the world, we must devise a set of frameworks to address the challenges inhibiting the growth of social movements then they will be able to help us understand and develop social movements in a bigger picture. Although this work has been motivated by an engineering diversity campaign, *ILookLikeAnEngineer*, by examining and studying this campaign in terms of user participation and messaging, I will propose two analytical frameworks that address the challenges uncovered in the next section which could be applied in other hashtag driven campaigns as well.

1.2 Hashtag Activism Campaigns Challenges

The three main components of a social movements are 1) resource mobilization, 2) activities and 3) leadership [4]. Here I will explain the challenges facing hashtag activism campaigns to become an efficient and real-time tool to spread the messages online.

1.2.1 Resource Mobilization

In hashtag activism, the primary resource is online social media. There are different types of social media platforms available nowadays such as Facebook, Twitter, Instagram, etc. Each has its own purpose but all of them share a common property that enables individuals to have their own voice and interact with other like-minded people, then eventually they can start their own online movements. One of the biggest issue with these platforms is tremendous amount of data generation and constant change in the trends and issues everyday. In Figure 1.2 [5], there is a huge leap between 2016-2020 in internet-connected devices as the same growth can be seen in global data generation in 2020. Analyzing the content manually in these platforms is extremely time consuming and laborious or almost impossible. This information overload demands some sort of automation to be created to help with social movement analysis. Although there are solutions to address this like hiring analyst like Mechanical Turks to help with the analysis, the real-time analysis where temporal anlaysis becomes important in development of the movements cannot be done in such regard. So there should be an approach to deal with these. By utilizing machine learning algorithms developed in artificial intelligence field, we can devise some tools and

[4]The details about each will be coverd in the next chapter, "background and related works"

[5]Extracted from a Financial Times report

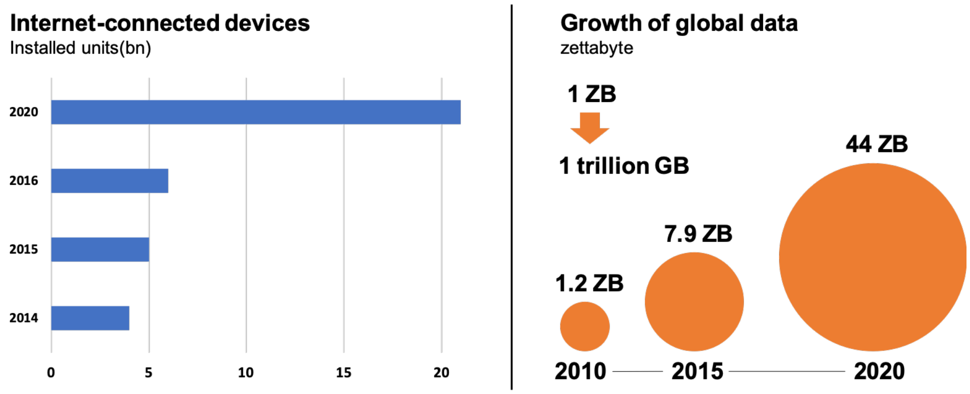

Figure 1.2: The trend in internet-connected devices and data generation

frameworks that can replace manual analysis to address most of the issues with high precision,or ("*Machine Learning Algorithms*").

Another issue that these online social media platforms have is a great deal of a great deal of concern in privacy and monetary gain as they are offered for free. So most of social media platforms have restricted the public access to the data on the platform to the point that it just allows to crawl his own personal information [6] or with limitations imposed in terms of time and count thresholds to download the data [7]. Sometimes to circumvent these restrictions is almost impossible or expensive. This makes analysis of a online activism impractical and economically infeasible as most of activisms are mostly launched by the marginalized and underrepresented minorities. Hence, financial viability for social movement where individuals are most likely not funded to promote their cause efficiently becomes a core of my work. To address this issue is to propose models that require no extra information with high accuracy, or ("*Real-Time Analysis*").

[6] `https://developers.facebook.com/docs/graph-api/overview/rate-limiting/`
[7] `https://developer.twitter.com/en/docs/basics/rate-limiting.html`

1.2.2 Activities

The convenience associated with online activism, e.g., distributing and signing petitions without anyone going door-to-door or just retweeting or writing a tweet with a hashtag, has generated such derogatory labeling of Internet-based collective action as "*clicktivism*" [Karpf, 2010,White, 2010], "slacktivism" [Christensen, 2011], or "engaged passivity" [Papacharissi and de Fatima Oliveira, 2012]. These names highlight a lack of commitment to the cause and limited effort and risk-taking on the part of members. Unlike activists that physically take part in protest marches or sit-ins, slacktivists do not put their bodies at risk [Stewart and Schultze, 2019]. By creating attractive, relatable and relevant content, we can lower the slacktivism in a given activism and inspire people to commit more to the ultimate goal of the movement, or ("*Message Analysis*").

Online activism provides an opportunity for people to connect with others who are taking public action or contributing to a common good to feel included in the cause, it is called *"connective action"* [Bennett and Segerberg, 2012]. For example, by using a hashtag, movement supporters can experience the feeling of being "part of something larger", getting "a sense of being there" –without being physically present –and sensing that "the whole world is watching –when physically present at a street demonstration" [Hopke, 2015]. This shows that social media movements with hashtag construct an individual's identity as an activist and raise the consciousness of individual members through the sharing of personal –rather than shared –experiences which is called *"collective solidarity"* [Radley and Bell, 2007], or ("*Use of Hashtag*").

1.2.3 Leadership

Leaders, influential actors, or movement entrepreneurs in an online social movements play an important role in shaping and directing the course of the movement. To communicate clearly about their messages, they are expected to convey to their desired audience specially when they are organizing collective actions. For example, Cardoso et al. (2019) [Cardoso et al., 2019] studied a digital activism about public garbage collection where the mobilization of participants was not clearly managed and left the different sets of group confused about the agenda. It showed that participants needed more than that, they needed to have a clear idea of how the collective action was going to work, what they were contributing to. To get around this in this example, other ICTs were more helpful, such as video sharing and blog (the how to), the waste map (the what), and the community website (to complement personal contact with the organizers and with friends and acquaintances involved).

Due to an extreme overflow of messages, the main issue was that messages needed to be much stronger and clearly targeted to the specific actors. Hence, understanding who is participating and contributing to the movement can help message creation become easier and targeted to the specific audience. Also finding out what kind of data each user type uses for their online presence will help us tune the messaging for each user type accordingly, or ("*User Analysis*").

These challenges with their respective solutions are summarized in Table 1.1.

Table 1.1: List of challenges and solutions for hashtag activism campaigns

	Challenges	**Solutions**
Resource Mobilization Online Social Media	• Dynamic nature and constantly changing. • Information overload.	• Machine Learning Algorithm. • Real-Time Analysis: no extra information required.
Activities Writing and sharing a tweet	• Clicktivism, slacktivism or engaged passivity • Lack of collective identity and solidarity	• Use of hashtag for SM (Hashtag Activism) • Message Analysis: understanding what is resonating with users.
Leadership Movement Entrepreneurs	• The mass of personalized messages. (the clear messages for collective actions are hard to find.)	• User Analysis: understanding who is participating and what kind of data is correlated to each user type.

1.3 Research Questions

The overarching issue framing this research project to address all of the aforementioned challenges in hashtag activism campaigns is: *With the data collection constraints and fast paced content generation on online social media platforms, how can we improve the success of a hashtag activism by generating more engaging content for specific user type in real time?* This question is broken down to two research questions as shown in Table 1.2.

Table 1.2: List of research questions with their respective outcomes

Research Question	**Analysis**	**Outcomes** (***Solution Identified***)
What is the user type distribution participating in the campaign in real-time?	**Real-Time User Type Classification**	• Identifying user type as individual(female/male) and organization in a unified real-time framework. (*Use of Hashtag, Machine Learning Algorithm, Real-Time Analysis*) • Understanding which feature of the message and user is more helpful with the task. (*User Analysis*) • Analyzing the user types in the ILLAE campaign. (*User Analysis*)
Is a tweet going to be engaging enough to help spread the campaign's message in real-time?	**Real-Time Retweetability Classification**	• Determining if a message based on the content and user information is going to be retweeted. (*Use of Hashtag, Machine Learning Algorithm, Real-Time Analysis*) • Understanding which set of features of the message and user is more helpful with the task. (*Message Analysis*) • Gaining insight into which features in the content of the message are more likely to engage more people in two hashtag campaigns. (*Message Analysis*)

1.3.1 Research Question 1: Real-Time User Type Classification

One of the main components of social movements are actors participating in a shared cause and one of the key issues for an online social campaign is to understand who is participating - specifically, whether the participants are individuals or organizations, and in case of individuals, whether they are male or female. This information can then assist in engaging those that seem disengaged or better leveraging existing participants to expand a campaign. Especially for campaigns that are looking to improve diversity, gender distribution and messaging across gender types is critical. Also classifying user type is important for identifying key actors in sustaining social movement. One of the main drawbacks in prior works on user type categorization is that they are computationally expensive due to the high dimensionality of feature representation extracted from textual data [Bergsma et al., 2013, Burger et al., 2011]. The sparsity of available information with a user profile has led to dependence on historical information such as the text of Twitter messages (tweets) from a user profile [McCorriston et al., 2015] for feature extraction, keywords in tweets [Volkova and Yarowsky, 2014], and user mentions in tweets [Bergsma et al., 2013] and friends/followers information [Chen et al., 2015]. Furthermore, all of the prior works have modeled the user type categorization problem as a binary classification (organization/individual or female/male) task and they have not explored any differences in the characteristics of user types that may be specific to the participation in a campaign community, in contrast to general user sample on a social network.

I devise a generic framework that classifies a Twitter user into three classes: organization, male and female in a real-time manner. A classification model based on the proposed framework utilizes different sets of features from a user. The proposed model can be used to analyze user type distribution of social campaigns to get real-time insights on the macro-level social behavior of participation and collective action. Moreover being able to classify the user type gives power to a social campaign moderator or influential actors to see who is engaging and participating the cause and adjust and tune the messaging or policy making.

The proposed framework uses just the information available in one tweet including user's data and one tweet data and classifies the user type into male, female or organization. The constraint for this classification problem is to use only one tweet from the user and further data collection will be avoided. Hence it meets the real-time requirement for the framework.

Later, this framework will be used to analyze the ILLAE campaign and will help us understand which feature sets defined in the framework are relevant to determining the user type. This gives us an insight into what entities are willing to participate in an online activism in this case, engineering diversity activism.

1.3.2 Research Question 2: Real-Time Retweetability Classification

One of the approaches to evaluate the level of social engagement in an online social campaign is retweetability analysis. One of the interesting emergent behaviors in Twitter is the practice of retweeting, which is the relaying of a tweet that has been written by another Twitter user. When a user finds an interesting tweet written by another user and wants to share it with his followers, he can retweet the tweet. Retweeting can be understood as a form of information diffusion since the original tweet is propagated to a new set of audiences, namely the followers of the retweeter [Suh et al., 2010].

In this work, a specific framework to hashtag activism campaign will be proposed to help a person writing a tweet to know if the tweet will be socially engaging or not (if it is going to be retweeted or not?). I hypothesize that using the cluster of various co-occurring hashtags shared in a social activism campaign helps people from other similar movements to connect where people can share the lessons learned through this exchange. Therefore, I devise another analytical framework that can predict if a given tweet will be engaging and attractive enough to be retweeted by other users based on a history of the messaging in the campaign. This will help activists and leaders of the campaign come up with appealing content and sets of more relevant hashtags to help disseminate the message. Therefore, this framework focuses on the features available in the content of one tweet and determines if it will be retweeted or not based on the history of the messaging in the campaign.

Later the factors affecting retweetability will be reviewed, given the fact that there are two main sets of factors [Chung, 2017], that could contribute to the retweetability; content (text, URL, photo and video) and user's social network properties [Pervin et al., 2014, Neppalli et al., 2016]. The next step will be that this analytical framework will analyze two different hashtag activism campaigns in terms of the subtopics/themes to see what themes or content of the text are good predictors of the retweetability. Knowing such information will benefit those who would like to engage more people in their cause.

1.4 Purpose of Study

The purpose of this work is based on how to raise awareness through online social movement for the family, students and female engineer to create a support system to keep them motivated in these fields. This addresses the problem of the often low retention of women in educational programs, engineering academia as well as the engineering profession. By devising a set of analytical frameworks, people are equipped better with how to compose a message for a specific audience on a hashtag activism. This way the message propagation will be more economically efficient and faster.

1.5 Significance of Studies

This research will contribute to the literature in the following ways. The fundamental contribution of this research is in shedding light on issues that have the potential to impact multiple stakeholders and how they pursue STEM related efforts. The two analytical frameworks provide an economically viable tools for those educating people through social media to be more efficient. First, the real-time user type classification will provide a tool for people specially influential actors or movement entrepreneurs to see the distribution of user type in a given campaign. Second, it helps an integrated framework to classify both individual and organization together, which is more important for online social campaigns where more organizations tend to engage than general setting. Third, this work will propose a framework for retweetability classification that helps a writer of a tweet know beforehand if that specific tweet will be retweeted or not. This helps him write a more engaging and attractive tweet for that given hashtag campaign. Fourth it will look more deeply into the factors affecting the retweetability task in terms of the tweet content and user's social network data to see if there is a high correlation between the content and the retweetability as a measure of information diffusion. Fifth, it helps us understand to know how to write an engaging tweet as an only option for the contributors in real-time for a specific campaign.

1.6 Summary

The core objective of this dissertation is to understand and address the aforementioned issues in engineering diversity related hashtag activisms. This work will help the literature with two real-time frameworks that can be extended and utilized in other hashtag campaigns as well. It does help with user type analysis in a campaign in real-time to determine what are the distribution of the contributors, hence the messaging will be adjusted. The second framework provides a tool for a content creation for a given hashtag activism campaigns. This will help those who are interested in raising others' awareness specially in this case, engineering diversity, know how to create attractive and engaging tweets. This dissertation is divided into 5 chapters. The first chapter has been dedicated to explain what is the focus of the research and why it is important to explore it further. It has concluded proposing two analytical frameworks to address the challenges discussed earlier. Chapter 2 will study the background of diversity in engineering and social movement with its important components from which the challenges were identified, then the main case study of this dissertation, *ILookLikeAnEngineer* campaign, will be explained. Chapter 3 will be discussing the details of the first analytical framework, the real-time user type classification. Then with the help of that, *ILLAE* campaign will be analyzed in terms of user type distribution and make some interesting observations. Chapter 4 will be explaining the details of the second analytical framework, the real-time retweetability classification, as an efficient tool for content creation for a given hashtag activism campaign. Later, the factors and features which are more likely to attract people will be studied. Finally, the dissertation will conclude in chapter 5 summarizing the contributions along with limitations, future works and applications of this work.

Chapter 2: Background and Related Works

This chapter studies and reviews the related works for the foundation of this dissertation. It starts with diversity in engineering as a major issue in STEM fields in both education and industry, then it presents the solutions to that. Later, social movement will be studied as an avenue to bring a social or political change to the existing issues specially engineering diversity issue here. Then it will study its fundamental components. The next section will look into hashtag activism campaigns as a new form of online social movement and it will be compared to the conventional social movement. This chapter will be concluded with a hashtag activism campaign case study for engineering diversity as the main motivation of this dissertation.

2.1 Diversity in Engineering

Diversity has become an important issue in globalization and competitiveness of institutions and corporate entities [Boyce et al., 2002, Palermo, 2004] . A Diverse workplace is usually defined as where "everyone feels included, appreciated and valued" and "everyone can achieve their potential" [Lewis et al., 2007]. But some professions like engineering may miss out on valuable contributions and new thinking approaches which a diverse workforce may add and which may indeed be needed to achieve the exceptional performance required to solve today's complex engineering problems. Lewis et al. (2007) [Lewis et al., 2007] have addressed this issue comprehensively in their diversity guide for the engineering profession.

Diversity has a number of dimensions ranging from personality to organizational contexts, with gender, race and ethnicity, physical ability and age probably being the most predominant differences, and, applied to engineering, the conflict between internal and projected dimensions may explain some of the marginalization effects observed [Waller, 2003]. To address the diversity challenges, the focus of this study is gender. The methodologies presented in this work, however, can be applicable to directly other diversity issues as well.

Diversity of the student body at academic institutions has immediate economic benefits [Schäfer, 2006]. In an age where many academic institutions compete for enrollments, in particularly in engineering, not drawing on the pool of women, for example, reduces revenue. This has motivated many institutions to maintain a level of recruitment and marketing activities ranging from open days, school visits, sometimes with special emphasis on women, to maintaining some form of women in engineering activity. "A department's willingness to accept responsibility for maintaining a student-initiated program is an obvious indicator of its stance toward women" [Kameda, 2011].

The image of engineering is masculine - indeed which is a huge image problem [Tietjen, 2004]. The perception of engineering as masculine is, according to Phipps (2002) [Phipps, 2002] a reason for the female minority. Hence the male over-representation and perception both cause and are caused by the underrepresentation of women. To achieve change, it is not new advantages or special measures that need to be created for women, in fact many of those create further resentment in men. It is men that really need to be convinced of the benefits of and the need for change.

While in Australia 15% of the undergraduate engineering enrollments are women, in the workforce this reduces to 5% [Roberts and Ayre, 2002]. Low retention has been attributed to "the unfriendly, even hostile culture of engineering" and the discrepancy between the need of women to "work cooperatively, and with a holistic perspective" as opposed to the more established hierarchical arrangements. In academia the "difficulty in collaborating with colleagues" leads to job offer rejections, marginalization of women in engineering positions and retention generally tends to be very low [Boyce et al., 2002].

Overall, educating men about the benefits of the diversity will alleviate the path toward a more friendly acceptance of women in STEM. Also educational programs at student level much depend on learning style and curriculum content where practices that enhance self-esteem and confidence are seen as valuable [Williams et al., 2002] as well as other incentives to interest women in pursuing education and career in engineering [Malik et al., 2018].

2.2 Social Movement

McCarthy et al. (1977) [McCarthy and Zald, 1977] studied the different approaches of understanding of social movement phenomena during the past decade and they concluded that they were different in a number of respects specially in those of [Gurr, 2015], [Turner et al., 1957], and [Smelser, 1963]. But, most importantly, the common theme among all of social movements was described in [McCarthy and Zald, 1977] as follows:

Shared grievances and generalized beliefs (loose ideologies) about the causes and possible means of reducing grievances are important preconditions for the emergence of a social movement in a collectivity. (p 1214)

Later, Diani (1992) [Diani, 1992] proposed three shared criteria of what constitutes social movements as follows:

1. There is a network of informal interactions among a diverse group of individuals, groups and/or organizations.

2. This network is engaged in some form of a political or cultural conflict.

3. There is an accepted sense of shared collective identity among the network.

Then Tilly (2019) [Tilly, 2019], adding a dynamic element to this, argues that social movements, in addition to being a network of actors, are a series of contentious performances, displays and campaigns such as processions, rallies, demonstrations, petition drives, statements to media, and so on. James et al. (2014) [James and Van Seters, 2014] combine these elements together and argue that there has to be formation of a collective identity, development of a shared normative orientation, sharing of a concern for change of the status quo, and occurrence of moments of practical action that are at least subjectively connected together across time addressing this concern for change.

Overall, when we look at different definitions and criteria of social movements, there are 4 main shared components as following:

1. Shared Cause/Goal

2. Collectivity (Action & Solidarity)

3. Actors/Users/Audience Participation

4. Resource Mobilization

As far as actors are concerned in social movements (SMs), there are 3 levels of people participating in SMs [Braccini et al., 2019] (see Figure 2.1: left circles); The outermost ring includes the mass of sympathizers who contribute with votes and the political strength of

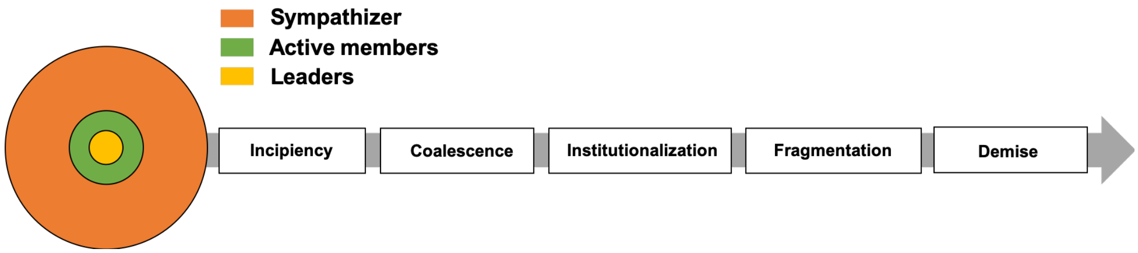

Figure 2.1: The 5 mains stages of a social movement

their numbers, the middle comprises a smaller number of active members committed to the movement's success and the innermost ring consists of formal leaders and coordinators.

Social movements have different stages of development as they grow larger. There are 5 stages for SMs [Braccini et al., 2019] (see right side of Figure 2.1) :

1. *Incipiency*: masses of individuals exchange their similar concerns, hold occasional meetings and communicate to build their momentum and shared identity.

2. *Coalescence*: SM starts adopting organizational structures and developing forms of formal and informal organization.

3. *Institutionalization*: the SM presents a society-wide organization, with a large number of members, sufficient resources, a division of labour and well-known spokespersons.

4. *Fragmentation*: the SM enters a divisive stage. People move away from the SM, either because they have been co-opted by other forces or they are pleased with the SM's battles won.

5. *Demise*: either the SM has achieved its goals, or all its resources have been acquired

by other forces.

2.2.1 Resource Mobilization

The resource mobilization theory, [McCarthy and Zald, 1977], is based on the notion that resources, such as time, money, organizational skills, and certain social or political opportunities, are critical to the success of social movements. Resource mobilization theory was the first to recognize the importance of influences outside the social movement. Also it has been criticized for its assumption of the constancy of discontent and collective interests over time, its overemphasis of the significance of outside resources, and its inability to adequately address social movements that begin with fairly substantial resources or those instigated by some minority groups. But when it was revisited later in 2011, it became an important element of recent social movement such as Egyptian Revolution [Eltantawy and Wiest, 2011].

The main idea of resource mobilization is to increase the reward and decrease the cost of joining the movement. It causes to increase the social motives for people which can lead to increasing user participation which improves the probability of the movement success.

2.2.2 Activities

Milbrath et al. (1977) [Milbrath and Goel, 1977] looked at the individual's level of commitment to a social movement and explained the political participation in terms of activities with increasing levels of effort and commitment, and to suggest that those at higher levels still engage in lower level undertakings and that political activity is cumulative [Huggins, 2002]. Starting with apathy (no political participation), Milbrath's hierarchy mapped out

Figure 2.2: Hierarchy of activities in traditional movements

thirteen political activities into three increasing levels of participation and decreasing levels of participants:

(1) spectator activities (the lowest level of commitment and effort and with the highest number of participants), (2) transitional activities (medium level), and (3) gladiatorial activities (highest level of effort and commitment and with the fewest participants). The hierarchy is illustrated in Figure 2.2.

Low level spectator examples include wearing political buttons or placing campaign signs in one's yard. Transitional activities require more effort, for example, attending a political rally. And the gladiatorial activities require significant commitment, such as organizing a political party. In traditional activism, spectator activities require the least effort, small amounts of time, insignificant amounts of money, and little thought or planning. Transitional activities are a greater burden to participants. Donations of money and time are found in this tier along with people attending meetings and contributing to their cause intellectually, physically, and financially. Last, gladiatorial activities command the greatest commitment.

2.2.3 Leadership

When a social movement is launched, everyone would like to think there is an organizer to initiate the movement [Schussman and Earl, 2004] . The characteristics of movement leadership were studied in the previous research, for example; a number of scholars have argued that significant activist biographies are important to movement leadership. In [Oberschall, 1973], regarding early social movement research, it was argued that professional and highly educated movement leaders are recruited to their positions via a history of political activism, student radicalism, and intellectual affinity to ideologies of protest or reform. It also argued that the entry costs of leadership are high because leaders –are exposed to considerably higher risks and penalties and have to expend more time, energy, personal resources, and make more economic sacrifices than the followers who subsequently join a social movement when its nucleus is already formed (p 159). That is why in most traditional movements the leaders are those with relevant experience on the topic which is usually associated with an increase in age.

2.3 Hashtag Activism Campaigns

In the previous section, the components of traditional social movements were studied and recently there is an increasing interest in studying them more [Rucht, 1991]. At the same time, citizens around the world have become increasingly aware of and interested in the expanding use of digital technologies –mobile phones and Internet-enabled devices, for example –in campaigns for social and political change [Joyce, 2010]. Hence the term, "Digital Activism" has been used by journalists, scholars, students, activists, and enthusiasts who are interested in studying the intersection of digital world with social movements.

The creation of digital technologies has changed the landscape of activism and digital activism - the use of various forms of digital technologies (e.g. blogs, videos, podcasts, email, and social media) for activism purposes - has gained significant momentum in recent years. Digital technologies enable activists in disseminating information quickly to a larger audience through a multitude of channels for a social, political, economic, or environmental change/cause [Joyce, 2010, Sandoval-Almazan and Gil-Garcia, 2014]. Activists use networked technologies not only for creating and sharing information but for forming public opinion, planning and calling for action, protect activists, as well as mobilizing both online and offline resources [Bastos et al., 2014, Borge-Holthoefer et al., 2015, De Choudhury et al., 2016, Segerberg and Bennett, 2011, Vegh, 2013].

Similar to offline activism campaign, most of the digital activism campaigns or cyber activism are either reactive or call for a proactive action [McCaughey and Ayers, 2013]. Reactive campaigning (e.g. regime change, government accountability, civil disobedience) is often politically motivated and targeted against certain controls or the authorities imposing those controls, meanwhile the proactive campaigns (e.g. disaster resilience, human rights, anti-bullying, and racial equality) promote or highlight a cause with an aim for achieving an certain objective [Bastos et al., 2014, De Choudhury et al., 2016, Sandoval-Almazan and Gil-Garcia, 2014, Vegh, 2013]. Digital activism has not only been used in developed countries but increasingly in the developing ones for resistance acts and organizing public protests [Bastos et al., 2014, Borge-Holthoefer et al., 2015, Eltantawy and

Wiest, 2011, He et al., 2015, Khondker, 2011]. Contrary to offline activism campaigns, digital activism campaigns are not bound by time or place. Online users can participate and engage without these restrictions to support the cause or get involved in registering their protest [Sandoval-Almazan and Gil-Garcia, 2014].

Within the realm of digital activism, Internet-focused activism have become increasingly common and the rise of social media has provided a new venue for networks to come and act together [Van Laer and Van Aelst, 2010, Taylor et al., 2004]. These platforms have allowed activists to cross borders of all kinds to connect, share and organize specially for women's rights despite national and cultural divides [Higgs, 2015]. One of the most interesting developments in digital activism in recent years is the rise of hashtag activism when large numbers of postings appear on social media under a common hashtag word with a social or political theme [Gunn, 2015, Yang, 2016]. Hashtags were initially introduced by Twitter to classify tweets into common theme or topics to facilitate easy search of specific messages and have been adopted by many other social media platforms including Facebook and Instagram [Van Dijck, 2013]. Hashtags have evolved over time and they are not only used for categorizing content but tailored and crafted by users for various purposes such as events, branding, breaking news, and supporting causes. In recent years, social media in general, and Twitter in particular has seen a surge of usage and is commonly associated with furthering a number of social and political issues [Bastos et al., 2014, De Choudhury et al., 2016, Eltantawy and Wiest, 2011, He et al., 2015, Sandoval-Almazan and Gil-Garcia, 2014, Tan et al., 2013], e.g. hashtags such as *#BlackLivesMatter*, *#YesAllWomen*, *#OccupyEveryWhere*, and *#BringBackOurGirls* are regarded as some of influential cases of hashtag activism [Yang, 2016].

"Hashtag feminism", or feminist activism has played a central role to inspire women

Table 2.1: Three first stages of a digital activisms

SM Stages	Incipiency	Coalescence	Institutionalization
Main Logic	Connective Action	Collective Action	Controlled Collective Action
SM Organizational Structures	Charismatic Leadership	Fomalized Rules & Roles	Formal leadership
Use of Digital Technologies	Freedom of use + Discussion Initiation	Select valuable content + content producers	Select valuable content + Educate valuable content producers

to fight gender inequities around the world by sharing their experiences and educating people to help with creating a momentum from online personal expressions to online collective action [Clark, 2016]. *Online social movement*, *online campaign or activism*, *hashtag activism*, *cyberactivism* and *e-movements* are interchangeably used in the research literature - as it has been the case when it has been referred to by its supporters and the media.

In Table 2.1, you will see the first three stages of digital activism [Braccini et al., 2019]. The differences between traditional SM and digital SM will be explained in the following subsections.

2.3.1 Resource Mobilization

Incorporation of social media as an important resource for collective action and the organization of contemporary social movements is the core of digital activism. Also digital activism needs social media as a key tool to form and publicize its identity. As we see in Table 2.1, there is a difference between the first stage, *Incipiency*, and the second stage, *Coalescence*, on *Main Logic*, as the the digital activism starts with *Connective Action* then it turns into *Collective Action* later in the second stage. Collective action is usually characterized by unstructured actions and interactions among members of a network, often occurring without central control [Wasko et al., 2005,Wasko et al., 2004,Nan and Lu, 2014] whereas connective action is a more basic step before the SM is shaped on online social media completely and it is more loosely committing the actors to the values and goals of the SM. Connective action is similar to collective action in that it involves individuals coming together, but connective action purposefully utilizes IS (Information Systems) to organize and communicate and often includes the use of social media [George and Leidner, 2019]. Also connective action focuses mainly on the activities performed by members of an SM to attract and mobilize new members in the digital world [Bennett and Segerberg, 2012].

One difference between collective and connective action is in how participants align with the values of a social movement. In traditional collective actions, participants are nearly always aligned with the ideas of the social movement. However, connective action does not demonstrate this trait and participants engage with varying levels of commitment and belief with the use of IT.

With the introduction of social media platforms, resource mobilization for digital activism has become an easily accessible resource where everybody with a little bit of IT knowledge can utilize and participate in their desired channels of communications. For instance, social media in the Egyptian revolution is how it changed the dynamics of social mobilization. Social media introduced speed and interactivity that were lacking in the traditional mobilization techniques, which generally include the use of leaflets, posters, and faxes [Eltantawy and Wiest, 2011].

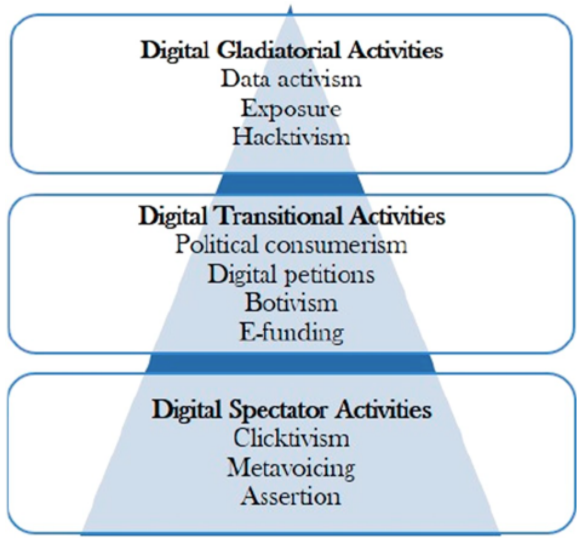

Figure 2.3: Hierarchy of activities in digital activism

2.3.2 Activities

George et al. (2019) [George and Leidner, 2019] used Milbrath's hierarchy of activities, see Figure 2.2, to come up with the equivalent for digital activism, see Figure 2.3. In traditional SM activities, Figure 2.2, you will see that spectator activities provide little impact while gladiatorial activities provide the most impact. But this is not so for digital activism, where major impacts can be found in every tier. For example, a hacker can create chaos with surprisingly little effort, and retweeting the time and location of a demonstration can result in a million protesters. Another point is that Milbrath's hierarchy increases with commitment and effort while digital activism hierarchy tiers vary according to the digital resources available. Digital resources include technical skills, technology artifacts, social networks, internet and communications access here [Selander and Jarvenpaa, 2016].

Digital Spectator Activities are the focus of my work and they are as follows:

1. **Clicktivism**: *Clicktivism* or *slacktivism* is “liking”, “upvoting”, or “following” an activist social media post or blog. While clicktivism indicates advocacy, it does not provide a voice to the participant to express original views. It is the easiest activity in the hierarchy.

2. **Metavoicing**: is *sharing*, *retweeting*, *reposting*, and *commenting* on a social media post created by another. It requires a little bit more dedication and interest from the user. This is where my research fits in.

3. **Assertion**: This action describes social media content creation. It informs others via video, audio, image, or text media. This is the highest level of participation in terms of interest and workload.

2.3.3 Leadership

Traditional movements have formal, hierarchical “closed-access” leadership allowing leaders to make decisions with relative autonomy. In contrast, digital activisms tend to depart from such centralized decision making in favor of more consensus-based decision making.

Social media has introduced a novel resource that provided swiftness in receiving and disseminating information; helped to build and strengthen ties among activists; and increased interaction among them and attracts the rest of the world’s attention. These platforms may be seen as an important, instrumental resource for collective action and social change empowering individual activists with sufficient knowledge of social media resources who can help to bring the digital movements to life without a lot of social or political movement experience. In another word, IT enables technical people with little political or social knowledge to become the leaders of their movement as opposed to online activism where leaders are expected to have deep knowledge of their cause and strong leadership background. These people are called *Movement Entrepreneurs*.

2.4 Activism - Then and Now

Table 2.2: List of differences betweetn *Traditional Activism* vs *Social Media Driven Hashtag Activism Campaigns*

Differences	Traditional Activism	Social Media Driven Hashtag Activism Campaigns
Resource Mobilization	leaflets, posters, faxes, events, ...	Social media platforms like Facebook or Twitter
Number of Participants	Successful social movements were associated with large numbers of participants.	Social media platforms provide efficiencies that allow fewer participants to have a greater impact.
Age of Participants	Greater participation was associated with an increase in age.	Younger people with social media skills are more likely to engage.
Leadership	Individuals with political or social movement experience or knowledge.	Anyone with good social media skills but political or social movement knowledge is not necessarily required (*Movement Entrepreneurs*).
Success Factors	An identified cause or campaign, effort, worthiness, unity, number of participants, commitment, resources.	Social medial skills, access to the internet, large social network.
Participants' Connection	Attending meetings and demonstrations, communications via post (mail), manned information tables.	Via social media through tagging relevant hashtags.
Participants' Activities	Political discussion, wearing a symbol, ...	Liking, writing or retweeting a tweet, ...
Marginalized Groups	Marginalized groups were often left on the sidelines because of a lack of resources.	Marginalized groups have more options to make their voices heard.

There are a number of differences between traditional activism and hashtag activism campaigns[George and Leidner, 2019] . Prior activism relied greatly on the number of participants [Tilly, 2006], while a smaller number of digital activists can create a substantial impact through the efficiencies afforded by technology. Participant age is another difference. Formerly, age increased as one increased their level of participation, with few young people running for office or organizing campaigns. The opposite is seen today as younger people tend to have the greater technical skills demanded at the higher levels of the digital activism [Rainie et al., 2012]. The opportunity for marginalized people to be heard is greatly expanded in digital activism compared to traditional activism [Schradie, 2018, Young, 2018]. As discussed before, leadership role has changed dramatically from traditional activism to digital activism by virtually relying on social media skills. This has created many opportunities for many with these skills to initiate their own movements [Earl and Schussman, 2002]. Traditional movements required a lot of commitment and resource allocation from participants such as time, creating banner, organizing an event but through social media with such as retweeting a tweet, people are notified of the movement much faster. These differences are summarized in Table 2.2.

2.5 #ILookLikeAnEngineer Campaign

This dissertation has been motivated by an engineering diversity hashtag activism campaign called *#ILookLikeAnEngineer*. This campaign is a good example of educating and empowering women in engineering field that helps with the stereotyping issue discussed before. Hence, this work has been formed based on this campaign and it will be further analyzed through answering research questions.

In the U.S., starting with the suffrage movement, a history can be traced for women's right that moves through the 60s and 70s leading to the 90s "new social movements" and "post-feminism". Within the "social movements" literature, women's right and feminism have garnered significant research and many of the theoretical developments have relied on studying these movements to further our understanding [Rupp and Taylor, 1999, Taylor et al., 1995]. In spite a long trajectory of social movements aimed at improving women's rights, there has been limited impact on women's role in broadening participation in the workforce, especially in Science, Technology, Engineering, and Mathematics (STEM) fields. The #ILookLikeAnEngineer therefore commonly shares issues of equality, diversity, and justice with prior efforts and displays many elements that constitute social movement in itself: it relies on informal interactions among supporters, the actors who participate are engaged in a common social justice cause; and there is a sense of a commonly shared identity

The #ILookLikeAnEngineer Twitter hashtag was an outgrowth of an advertising campaign by the company OneLogin. In late July 2015 the company OneLogin posted billboards across public transport in the California Bay Area, especially at the BART train stations, as a recruitment campaign. The billboards depicted different engineers working in the company at that time. In this series of billboards, the engineers all wore black t-shirts and were photographed from their head till about their waist. The billboards also exhibited the engineer's name, job description, and a very short quote alongside the photo. One of the billboards depicted a woman engineer named Isis Wenger (now Isis Anchalee and that is how we refer to her in the paper) and her photo attracted a lot of attention on the Web (Figure 2.4).

Her image led to discussions online about the veracity of the campaign as some people found it unlikely that she was really an engineer. Online comments stated that she was

Figure 2.4: Billboard campaign featuring *Isis Anchalee*

"too attractive" to be a "real engineer", among other demeaning comments. These online discussions, according to Anchalee, prompted her to write a post on Medium, on August 1, 2015, responding to the stereotypical reaction her image in the ad had generated. She stated in the post that, "At the end of the day, this is just an ad campaign and it is targeted at engineers. This is not intended to be marketed towards any specific gender –segregated thoughts like that continue to perpetuate sexist thought patterns in this industry." On August 3, 2015, in addition to a Tweet she updated her initial blog post to add a call for action and an image with a Twitter hashtag: Do you feel passionately about helping spread awareness and increase tech diversity? Do you not fit the "cookiecutter mold" of what people believe engineers "should look like?" If you answered yes to any of these questions I invite you to help spread the word and help us redefine "what an engineer should look like #ILookLikeAnEngineer" (see Figure 2.5).

She wanted to show that she was not the only female engineer and also take some of the attention away from her and towards the issue of diversity in engineering. "When this is all in the past, I would like to have created a larger impact on the community than

Figure 2.5: Start of ILookLikeAnEngineer hashtag activism campaign on Twitter

simply generating a Twitter presence. Spreading awareness is the first step, but I want to help facilitate concise plans of action so we can create a genuine change. If you have any personal input and advice that you think could help make a difference please feel free to share it with me. You can help support us by checking out [*link to a new billboard campaign followed by link to an event*]."

Her Medium post as well the hashtag received large support and the Twitter tag in particular took off as a way for others to share their images and their ideas on the issue. The hashtag soon saw significant media coverage not just in the US but also across the Atlantic, particularly in the UK. The movement not only had an online presence but a large event was organized in September 2015 in the Bay Area supported by a group. The hashtag movement was started as a campaign to raise funds to be able to post billboards across the Bay Area with the hashtag and photos; in effect, to take the online movement into the offline world through a series of billboards. Overall, the use of the Twitter hashtag, combined with media coverage and the decision to post billboards across Silicon Valley turned the effort into a larger scale activism effort, in essence, a movement for recognizing women's participation in engineering.

Chapter 3: Real-Time User Type Classification

3.1 Introduction

In this chapter, I propose a specific way with the use of machine learning algorithms which can help those engaged in social media activism to better target their messaging and improve their engagement with participants. It is commonly reported that one of the primary deficiencies of social media campaigns is that they only engage those who are already interested in the topic of the campaign, often creating "filter bubbles". The other criticism targeted towards activism campaigns on social media is that they reflect a form of laziness by participants (*slacktivism*) and online participation does not translate into offline action. Many researchers and practitioners argue against this characterization of social media campaigns by arguing that there are residual effects of participating in these efforts and people who participate in social media activism campaigns are more likely to participate in other similar causes [Lee and Hsieh, 2013, Skoric, 2012]. There are other considerations of such campaigns as well, such as whether they are being co-opted by other actors - media, corporations - and to what extent.

As studied earlier, user participation is one of the principal elements of a social movement. And diverse group of users/organizations plays an inescapable role in bringing people together for the common goals. Although every online activism does not get molded into a movement, social technologies create easily accessible platforms to support diverse participation in social movements and allow users to engage with a social movement in ways that are different than participation in traditional social movements. In particular, users can now connect loosely – through signaling their support via a tweet, for instance – allowing for different forms of organizing or connective action. Hence, classifying user type is important for identifying key actors in sustaining social movement. In a recent article, [Vaast et al., 2017] suggest that different users might take on different roles within the activity, "complementary, interdependent roles that make up the connective action". These relationships, according to them, have not been investigated empirically and there is a need to better understand social media that enable connective action [Vaast et al., 2017]. Therefore, gaining insight into user demographics such as gender, age, ethnicity, etc. [Bergsma et al., 2013, Volkova et al., 2015] is essential for promoting wider social engagement. One of the most critical attributes among them is gender (user type) given the prevalent gender stereotyping in the society. [Johri et al., 2018b] studied #ILookLikeAnEngineer campaign in different aspects, one of which was "what are the triggers that provide momentum and sustain the campaign in the early stages?" It then identified 4 different triggers of activity to help the movement; 1) event-driven, 2) media-driven, 3) industry-driven and 4) personality-driven. These 4 triggers or driving forces for the social movement can be categorized into two entities: individuals and organization. By automating user type identification of the movement, we can have a more comprehensive analysis on the movement. Moreover, [Malik et al., 2018] studied this movement in terms of user participation in different regards. The user types of most retweets include, female, male, company, university and online community. As you can see, these two papers show that the presence of companies (non-individual entity) is prevalent in sustaining the movement along with individuals(female/male). Thus, the key objective of this section is to develop a method to automatically identify user types in real-time for supporting analysis of campaign dynamics, whether they are individuals or organizations [McCorriston et al., 2015, De Choudhury et al., 2012, De Silva and Riloff, 2014] as well as female/male [Burger et al., 2011, Chen et al., 2015, Bergsma et al., 2013].

3.2 Related Works

The rapid growth of social media in recent years, specially Facebook and Twitter, has led to a massive volume of user-generated content including informal text and multi media. This in turn has generated a great deal of research interest in aspects of social media, including automatically identifying latent demographic features of online users specifically gender(female/male) [Schler et al., 2006, Burger and Henderson, 2006, Argamon et al., 2007, Mukherjee and Liu, 2010].

Tweets are tagged with many sources of potentially discriminative metadata, including timestamps, user color preferences, icons, and images. In [Burger et al., 2011] the textual sources of features such as screen name, full name, description and tweet text are used for the classification task using both word and character-level ngrams from each field.

In [Chen et al., 2015], the interaction behaviors between Twitter users mainly include following and friending in addition to other more implicit ways like favoriting, retweeting and user mentioning. Also users' own posts or self-descriptions are not used as textual sources but only those generated from their neighbors. A collection of tweets/self-description aggregated from neighbors including followers and friends with up to 200 tweets per neighbor. To capture the lexical usage and topics, two types of linguistic features are used, n-grams and hidden topic distributions derived from a Latent Dirichlet Allocation (LDA) model. To enrich the feature set, profile images are used as well. The image features are based on the popular scale invariant feature transformation (SIFT) [Lowe, 2004].

On the other hand, the presence and activity of organizations on social media platforms has become interesting to be investigated further. Beyond corporate advertising and customer engagement initiatives, organizations like political parties, social groups, and local communities also use these platforms for communication and coordination [Golbeck et al., 2010]. The ability to distinguish between organizational and personal accounts can have significant advantages for applications such as election predictions [Tumasjan et al., 2010], health monitoring [Schwartz et al., 2013a], and crisis response [Saleem et al., 2014]. Despite known organization activity on social media and its affects on large-scale measurements of human behavior, little is known about the scale of this presence. In [McCorriston et al., 2015], a binary classification (personal or organizational) task is formulated with user's historical tweets using support vector machine algorithm.

One of the main drawbacks in prior works on user type categorization is that they are computationally expensive due to the high dimensionality of feature representation extracted from textual data [Bergsma et al., 2013, Burger et al., 2011]. And also the sparsity of available information with a user profile in real-time has led to dependence on historical information such as the text of Twitter messages (tweets) from a user profile [McCorriston et al., 2015] for feature extraction, keywords in tweets [Volkova and Yarowsky, 2014], user mentions in tweets [Bergsma et al., 2013], and friends/followers information [Chen et al., 2015]. Furthermore, all of the prior works have modeled the user type categorization problem as a binary classification (organization/individual or female/male) task and they have not explored any differences in the characteristics of user types that may be specific to the participation in a campaign community, in contrast to general user sample on a social network. The main 4 related papers are summarized in Table 3.1.

Table 3.1: List of main related works for real-time user type classification

Study	Limitations	Proposed Framework
[Burger et al., 2011]	• Not Unified User Type Framework • Not Computationally Efficient	• **Unified User Type Framework** (Individual (female/male) + Organization) • **Computationally Efficient** (Lower Feature Space) • **Real-Time** (no historical tweet or ego network, just one tweet)
[Bergsma et al., 2013]	• Non Unified User Type Framework • Not Real-Time	
[McCorriston et al., 2015]	• Not Unified User Type Framework • Not Real-Time	
[Chen et al., 2015]	• Not Unified User Type Framework • Not Real-Time	

3.3 Contributions

The main contribution in this framework is a method for assisting real-time characterization of users involved in a social media campaign, specifically the ILLAE campaign on

Twitter. Our findings suggest that female users are both more numerous and active than males. Unlike other social movements, this campaign attracted a much higher number of organizational participants and they are more engaged and tweet at a higher level. When combined with other information that is easily available, this characterization can provide information on which organizations and what messaging, etc.

Characterizing users has been attempted in many studies. Our aims in this work are to improve on prior work in two aspects:

1. Characterizing not just individuals into male/female but also into individuals and organizations.

2. Creating a process and framework that allows real-time analysis as opposed to using information and features that can only be assessed or computed retrospectively.

Overall, I propose a user type classification framework with the following contributions:

- It is a unified approach to address user type classification problem by combining all user types together, individuals (male/female) and organizations into one framework.

- Our framework addresses Twitter API limitations by using features available in real-time in the user's one tweet and user profile for the classification task as opposed to other methods which require additional information about the user or user's social network, such as friends, followers, and user mentions.

- The proposed framework encompasses features from multiple characteristics of information such as user profile's description, name, tweet and profile images making it comprehensive.

- Although automatic user type classification suffers from several caveats such as low recall [Minkus et al., 2015] and classification error [Yadav et al., 2014], it is very computationally and financially efficient to study online social media activism in a real-time manner.

3.4 Proposed Multimodal Analytics Framework

Here, I propose a framework with two different structures, hierarchical and flat, that uses multiple types of information from a user's one tweet and user profile information to classify the users into male, female or organization. This framework uses multi-modality of information from a user: screen name, image, metadata, and a tweet text. I use two types of structures, hierarchical and flat for this framework. For flat structure, Figure 3.1, the framework is a multi class classification that was not used in any previous research as all of the frameworks were binary classification. For hierarchical, Figure 3.2, the framework is a two step binary classification. The first step is a binary classification that classifies each user into individual and organization then for the second step, a gender database which

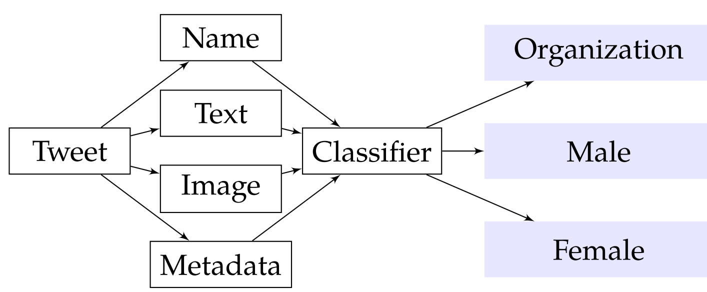

Figure 3.1: Proposed framework (**Flat Structure**) for user type classification with the help of multi-modality of information

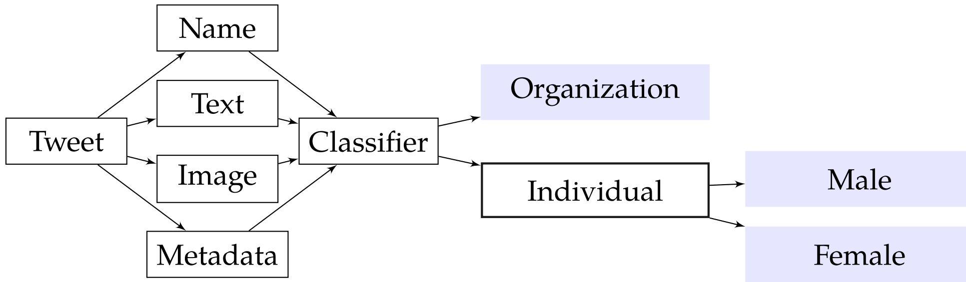

Figure 3.2: Proposed framework (**Hierarchical Structure**) for user type classification with the help of multi-modality of information

will be explained later, is used to determine male/female gender of the individual. In this section, I explain every feature of these two frameworks as both use the set of same features.

3.4.1 Name

Liu et al. (2013) [Liu and Ruths, 2013] showed that including name as a feature in existing classifiers yielded a 20% increase in accuracy. So it is a good indicator of individual gender, male/female. Besides, people participating in this kind of social movement, specially ILLAE where its sole purpose is about engineering diversity for women, don't usually hide their names and actually would like their voice heard. Therefore, name is a great discriminant feature for user type classification. I used a name database[1] that tries to infer a person's gender from their name. This database is created by the data that came from national statistics institutes and was accompanied by frequency information. A Twitter user has a screen name and Twitter handle. Although some Twitter handles has some information about the user, with preliminary experiments, the performance on screen name

[1] `https://github.com/tue-mdse/genderComputer`

is much higher than that of Twitter handle making screen name the input for the database that classifies users into female, male, unisex and none. So it is highly effective for individuals' name but for organizations, additional information is needed that will be introduced in the following sections.

3.4.2 Text

Understanding unstructured data, text, has been studied in different approaches. In prior works, character-based features generated by n-gram method (Bag of Word (*BOW*)) were used to encode the text. But the main drawback of this method is the lack of generalization to capture unseen input features because the model was unable to learn those features that were not created in the training phase [Bergsma et al., 2013]). To overcome this, I used an approach to analyze the language used in these texts by word-category lexicon; the most widely used is Linguistic Inquiry and Word Count (LIWC), developed over the last couple decades by human judges designating categories for common words [Pennebaker et al., 2015,Pennebaker et al., 2003,Schwartz et al., 2013b]. By LIWC, the psychological value of language in gender can be quantified with 15 different categories (93 subcategories) of language ranging from part-of-speech (i.e. articles, prepositions, past-tense verbs, numbers,...) to topical categories (i.e. family, cognitive mechanisms, affect, occupation, body,...), as well as a few other attributes such as total number of words used [Pennebaker et al., 2015,Chung and Pennebaker, 2007].

As ILLAE's collective action called for women working in engineering fields to post a selfie with #ILookLikeAnEngineer, using linguistic feature from LIWC can be a good predictor for the gender of the individual as female. Also in[Johri et al., 2018a], LIWC was used to filter out first-person tweets, then coded the content of those tweets into multiple categories. It found out the prevalent theme in these tweets from women expressing their engineering identity. Also females use more emotion words [Mulac et al., 1990, Thomson and Murachver, 2001], and first-person singulars, and they mention more psychological and social processes [Newman et al., 2008]. Males use more swear words, object references [Newman et al., 2008, Mulac et al., 1986]. Schwartz et al. (2013) [Schwartz et al., 2013b] shows that female are more agreeable and exhibit more psychological tone in their messages. All of this prove that using text with the help of LIWC helps with user type classification task.

LIWC2015 was applied to the textual part of the tweet and the profile description of the user and they were fed into the classifier as the text features.

3.4.3 Image

User profile images are very popular on social network and Twitter is not an exception. This gives an opportunity to leverage the visual information provided by users. The interesting property about them in an online social movement like ILLAE, as mentioned earlier, people usually have their own profile picture and use their authentic account to participate in this cause. So most of user profile pictures of individuals are their own head-shot, for example the first row of Figure 3.3. On the other hand, organizations most likely have non-human profile pictures, for example the second row of Figure 3.3. So using an image recognition framework helps us distinguish between individuals and non-individuals (organizations) better.

There are many algorithms such as SIFT [Lowe, 2004], SURF [Bay et al., 2006] etc. for image recognition task but due to handcrafted rules for feature extraction process in these algorithms, I decided to use one of the automatic feature extraction algorithms such

Figure 3.3: Few sample of user profile pictures, first row: *individuals* and second row: *organization*

as state-of-the-art Convolution Neural Networks (ConvNets). Among many ConvNets models, I selected VGG-16 model – a 16-layered convolutional architecture – due to fast computational capability [Simonyan and Zisserman, 2014]. The model takes a picture and returns 1,000 probabilities, each of which is a weight to a class trained by ImageNet[2] dataset.

VGG-16 trained with ImageNet dataset was used for user profile pictures to give 1000 numerical features for the framework.

3.4.4 Metadata

While content generated by users are important such as tweet, user profile description and profile image, the user's social network structure has also interesting features [Pervin et al., 2014,Neppalli et al., 2016]. Organizations have usually high follower and low friends count and have verified status. In Figure 3.4, you can see few examples of the organization. So these metadata help the framework identify organization better. The metadata of social network structure used in the frame work are followers count, friends count, tweets

[2] `http://www.image-net.org`

Figure 3.4: Few sample of organization account

(statuses) count and verified status.

3.4.5 Model Fitting

Based on the feature sets explained in the previous section, the following 4 categories of features are created:

- **Name**: Given user name, it predicts gender categories; female, male and unisex (3 categorical)
- **Text**: Given a tweet text and profile description, it gives you features of LIWC. Except for word count, others are the percentage of each category ($2 * 93 = 186$ numerical)
- **Image**: Given user profile image, it returns 1000 probabilities to which image class it belongs defined and trained in ImageNet (1000 numerical)
- **Metadata**: friends, followers and tweet counts with verified status (3 numerical and 1 categorical)

The expected output is the three classes: female, male, or organization.

The feature space for the classifier, 1193 features, is much fewer than that of other classifiers, e.g.: more than 15 million in [Burger et al., 2011] or 8^{3*5} in [Alowibdi et al., 2013].

3.5 Datasets

For this classification, two datasets of users from Twitter social network were used; one domain-specific (online social campaign – ILookLikeAnEngineer) and one general dataset (CrowdFlower[3]). Each dataset consists of the tweet metadata (such as creation time, tweet id, tweet text, location, user mentions and etc.) and user metadata (such as user's screen name, profile image, friends/followers/tweet count and etc.). The reason that the general dataset was selected was to test the generalization power of the framework and see if it is able to classify user types well. To test the stability of the framework in the training phase, the following datasets were used as original (imbalanced) and balanced in terms of training instances of each class. For making balanced dataset, random sampling was done from the two majority classes as many as the minority class.The results from these two kinds of dataset split showed the robustness of the framework against imbalance issue.

3.5.1 ILookLikeAnEngineer (ILLAE)

The domain-specific dataset was based on ILookLikeAnEngineer campaign, an engineering diversity campaign as explained in the introduction section. The dataset was first collected from Twitter using streaming API and based on three hashtags; **#ILookLikeAnEngineer**, **#LookLikeAnEngineer** and **#LookLikeEngineer**, as these hashtags were interchangeably used. The time frame for the data ranges from August 3rd, 2015, the day the hashtag was first used, until October 15th, 2015, which is about 2 months after the first

[3] https://www.kaggle.com/crowdflower/twitter-user-gender-classification

Table 3.2: **ILLAE** dataset with user distributions for imbalanced and balanced split

	Male	Female	Organization	Total
Imbalanced (Original)	353 *24.62%*	451 *31.45%*	630 *43.93%*	**1,434** *100%*
Balanced	353 *33.33%*	353 *33.33%*	353 *33.33%*	**1,059** *100%*

initial surge of the campaign. The dataset consists of 19,492 original tweets from 13,270 unique users (individuals and organizations). Three reviewers from our research team individually annotated 1,434 unique user profiles as female, male or organization. The annotated user profiles were reviewed by the team and in case of a disagreement, the user type was decided by mutual discussion. The distribution of each user type for imbalanced and balanced training instances for ILLAE is shown in Table 3.2.

3.5.2 CrowdFlower

This dataset was used to train a CrowdFlower AI gender predictor [4]. Contributors were asked to visit a Twitter profile link and judge whether the user was a male, a female, or a brand (non-individual). The dataset contains 20,000 user profiles, each with a username, a random tweet, account profile and image, location, and even link and sidebar color. Because of multiple annotation results by different people, each row has a certain degree of confidence signifying the likelihood of the majority gender. In order to make sure that I

[4] `https://www.crowdflower.com/using-machine-learning-to-predict-gender/`

Table 3.3: **CrowdFlower** dataset with user distributions for imbalanced and balanced split

	Male	Female	Organization	Total
Imbalanced (Original)	3,644 *36.31%*	3,939 *39.24%*	2,454 *24.45%*	**10,037** *100%*
Balanced	2,454 *33.33%*	2,454 *33.33%*	2,454 *33.33%*	**7,362** *100%*

worked on the cleanest data possible, I filtered every user type that had a degree of confidence less than 1. After filtering the dataset, there are 13,926 users. Because most of the URL of user's profile picture didn't work, I had to recollect the users' data from Twitter API based on their Twitter handle. At the end, I was able to retrieve 10,037 unique users. The distribution of each user type for imbalanced and balanced training instances for CrowdFlower is presented in Table 3.3.

3.6 Experiments

3.6.1 Preprocessing

Once all features are generated as described before, both log-count and normalization of each feature vector were used as preliminary experiments show the improvement compared to the original data. Some of the user profiles have no image or non-working image URL. Therefore, single imputation was used to replace the non-existing feature value for the respective user by using the average of existing values of that particular feature.

3.6.2 Support Vector Machines as Classifier

The features after preprocessing become ready to train the SVM classifier for the prediction task. The reason for employing an SVM classifier is due to SVMs becoming one of the most popular machine learning techniques and has been proven effective in non-linear feature space due to the kernel use. SVMs have shown great success in difficult classification problems and have a solid theoretical foundation in statistical learning theory [Vapnik, 1995, Vapnik and Vapnik, 1998]. The hyper-parameters for SVM used in the experiments are presented in Table 3.4.

Table 3.4: List of Hyper-Parameters of SVM used in the experiments

	Type	**Hyper-Parameter**
Algorithm	**SVM**	kernel=[linear, poly, rbf, sigmoid] cost = [0.01, 0.01, 0.1, 1, 10, 100, 100] gamma = [0.001, 0.01, 0.1, 1] coef0 = [0.1, 1, 10, 100] degree = [2, 3, 4]

3.6.3 Model Selection

There are different methods for model selection [5]. Hold-out cross validation was used for training and testing the framework and the baselines. 20% of the dataset was held out for the final testing and the rest of it (80%) was used for 10 fold cross validation. For

[5] For other methods, go to appendix in **Model Selection** section.

performance measure in hyper-parameter optimization for each CV phase, "accuracy" was used. Then based on the best set of hyper-parameters in cross validation phase, the training set was used for training then it was tested against the held out test set. "Accuracy" and "F1 score" for each class were used for performance measure of each model. To compare each classifier, I used McNemar's test [McNemar, 1947] which is a non-parametric test used to analyze matched pairs of data [6]. This test is better suited here instead of paired sign test or t-test as these both statistical tests have a prerequisite for the data that both of them should be independent from each other whereas this fails when comparing two classifiers trained from the same dataset. Mc Nemar's test doesn't have this prerequisite and it makes a better and robust statistical test for result comparison. The smaller the p-value, the stronger evidence it is that one classifier has better performance over another. The p-value threshold was set 0.05.

3.6.4 Baselines and Frameworks

As there two types of classification structures, flat and hierarchical, there are the following baselines along with the proposed framework:

- **Flat**: 3-class classification baselines with the proposed framework

    1. *Historical Tweets* (*Trained*): uses the feature modeling in [McCorriston et al., 2015] [7] .

    2. *Bag-Of-Words*

    3. ***Proposed Framework*** (***Flat***)

[6]For further explanation, go to appendix in **Statistical Comparison** section.

[7]`https://github.com/networkdynamics/humanizr`

- **Hierarchical**: 2 step classification as discussed earlier, first step as a binary classification (individual/organization) then using the gender database, male/female are determined.

  1. *Historical Tweets* (*Pretrained*): uses the pretrained model from the data in [McCorriston et al., 2015] [7]
  2. *Historical Tweets* (*Trained*): uses the feature modeling in [McCorriston et al., 2015] [7]
  3. ***Proposed Framework*** (***Hierarchical***)

Historical Tweets baselines has 110 features derived from last 200 tweets and user metadata that are trained by SVM. Even though it uses historical tweets from the users and it cannot be applied to real-time classification task, our proposed framework outperforms this as you can see in the following section.

3.6.5 Experimental Results

In this section, the experimental results of each dataset for imbalanced and balanced distribution will be analyzed and the best model will be used to analyze the entire hashtag activism campaign in terms of user type distribution.

The results for *imbalanced*(*original*) distribution of the datasets are presented in Tables 3.5, 3.6 and for *balanced* distribution of the datasets are shown in Tables 3.7, 3.8. The results are split based on their classification structure as discussed in the previous section.

The statistical comparison result with McNemar test were that both the frameworks (flat and hierarchical) against 4 baselines in both datasets for both *imbalanced*(*original*) and *balanced* were statistically significant ($p < 0.05$). Meanwhile, the proposed frameworks

Table 3.5: Results from **ILLAE** dataset with *imbalanced*(*original*) distribution - **6.65%** accuracy increase

Structure	**Frameworks**	Accuracy	F_1**-Org(%)**	F_1**-Female(%)**	F_1**-Male(%)**
Flat	Historical Tweets	70.77%	83.46%	68.37%	43.40%
	Bag-Of-Words	70.42%	83.10%	70.10%	31.11%
	Proposed Framework	**88.03%**	**95.16%**	84.82%	79.07%
Hierarchical	Historical Tweets (Pretrained)	81.34%	86.27%	82.93%	84.13%
	Historical Tweets (Trained)	77.82%	82.35%	83.44%	78.69%
	Proposed Framework	87.68%	94.35%	**88.10%**	**87.88%**

Table 3.6: Results from **CrowdFlower** dataset with *imbalanced*(*original*) distribution - **9.72%** accuracy increase

Structure	**Frameworks**	Accuracy	F_1**-Org(%)**	F_1**-Female(%)**	F_1**-Male(%)**
Flat	Historical Tweets	65.92%	69.75%	69.57%	59.03%
	Bag-Of-Words	50.92%	55.11%	52.33%	46.97%
	Proposed Framework	**75.64%**	**79.62%**	**76.65%**	**72.13%**
Hierarchical	Historical Tweets (Pretrained)	59.39%	69.50%	69.46%	70.89%
	Historical Tweets (Trained)	60.49%	73.81%	68.96%	70.13%
	Proposed Framework	73.13%	79.61%	70.16%	72.13%

Table 3.7: Results from **ILLAE** dataset with *balanced* distribution - **10.58%** accuracy increase

Structure	**Frameworks**	Accuracy	F_1-**Org**(%)	F_1-**Female**(%)	F_1-**Male**(%)
Flat	Historical Tweets	68.90%	80.28%	67.13%	58.65%
	Bag-Of-Words	63.16%	67.20%	66.67%	56.77%
	Proposed Framework	**87.08%**	**89.05%**	**84.85%**	**87.25%**
Hierarchical	Historical Tweets (Pretrained)	76.56%	80.82%	79.67%	78.79%
	Historical Tweets (Trained)	74.64%	76.81%	79.03%	81.20%
	Proposed Framework	81.82%	87.88%	81.89%	87.14%

Table 3.8: Results from **CrowdFlower** dataset with *balanced* distribution - **3.12%** accuracy increase

Structure	**Frameworks**	Accuracy	F_1-**Org**(%)	F_1-**Female**(%)	F_1-**Male**(%)
Flat	Historical Tweets	69.25%	76.31%	69.51%	60.60%
	Bag-Of-Words	51.26%	60.99%	48.72%	45.38%
	Proposed Framework	**72.37%**	80.51%	**70.34%**	66.73%
Hierarchical	Historical Tweets (Pretrained)	59.47%	72.83%	67.97%	67.93%
	Historical Tweets (Trained)	61.85%	78.46%	67.17%	65.28%
	Proposed Framework	71.02%	**81.82%**	67.33%	**68.26%**

with imbalanced and balanced distribution were not statistically significant. It shows that accuracy as a performance measure for model selection is not a choice and running statistical tests give a more plausible answer like McNemar here. But for the sake of result comparison, I evaluate the improvement of the proposed framework (flat) as the performance measure. That being said, in both cases (imbalanced and balanced) for both datasets, there is an increase in accuracy in proposed framework against the other 4 baselines. For *imbalanced*(*original*) distribution, 6.65% and 9.72% accuracy increased in ILLAE and CrowdFlower datasets respectively (Tables 3.5, 3.6). Also for *balanced* distribution, we can see accuracy increased 10.58% and 3.12% in ILLAE and CrowdFlower datasets respectively (Tables 3.7, 3.8). It shows that the proposed framework is robust to imbalance dataset to some degree and can be used both in general domain and hashtag activism campaign like ILLAE. An interesting observation here is the framework gives the highest accuracy increase in balanced and ILLAE dataset which shows that the feature selection in the previous section was very effective and relevant to the task as opposed the result of CrowdFlower for balanced dataset which is the lowest. It also reinforces the idea that the proposed framework works much better in the campaigns where people would like to genuinely make the change and their authenticity of their account is more prevalent than a general/random dataset. Another observation is the proposed framework outperformed the baselines with historical tweets. It shows that with a better feature selection from the existing data, the more economically efficient and accurate framework can classify the user types better.

The F1 score for each class is almost close to each other except in Bag-of-Words classifier in Table 3.5. In this case, the difference between the weakest class and the next one is huge (*31.11% for F1-Male to 70.10% F1-Female*). To figure out what caused this huge gap, the word distribution for each class was calculated. The mean and standard deviation

Table 3.9: User type distribution for cluster analysis in entire ILLAE dataset

	Male	Female	Organization
cluster 1	183	171	172
cluster 2	71	91	**376**
cluster 3	99	**189**	82

for each class is as follows: organization (30 ± 7), female (28 ± 9) and male (26 ± 9). All of these distribution were statistically different ($p<0.05$). It means that male users in the ILLAE campaign used fewer words to communicate, hence making it harder to be separable. To further the argument, a cluster analysis was conducted to see how each class was grouped. KMeans clustering algorithm was used with 3 clusters and each class distribution of clusters are presented in Table 3.9. It shows that in cluster 2 is dominated by organization and female user type has the majority in cluster 3, meanwhile in cluster 1, all user types have almost the same instances. The cluster analysis reinforces the argument that male users were harder to be classified compared to other two user types; organization and female.

To find out the user type distribution in the entire ILLAE dataset, the proposed framework (flat) was trained with the annotated profile users (1434) and then was applied to the rest of the users (11,863) to see the user type distribution in the entire dataset. For those users with missing value in their features (e.g.: image), single imputation is used to replace the non-existing feature value by using the average of that particular feature from annotated user profiles. The user distribution of the ILLAE dataset, Table 3.10, shows that female users are the most dominant user types in this online social campaign with 47.95%

Table 3.10: User type distribution in entire ILLAE dataset

	Male	Female	Organization	Total
User	2,222 *16.74%*	6,362 *47.95%*	4,686 *35.31%*	**13,270** *100%*

which is followed by organizations with 35.31%. Organization's presence in this dataset exceeding 35% is completely much more than that of a general/random dataset which is less than 10% [McCorriston et al., 2015]. This confirms the fact that in an online social campaign like ILLAE, organizations try to establish their presence to not only promote the cause but also create a new venue to spread their message and gain more marketing share. At last, male users did not engage in this social campaign as much as other user types as the primary audience target for this campaign was female.

On the same note, I would like to see the tweet, retweet and favorite distribution for each user types in the entire ILLAE dataset. In Table 3.11, female is the most dominant user type in all of the categories as it shows the importance of this campaign among women. Although the tweets created by female has the highest share, the percentages in retweeted and favorited tweets from female are higher. It shows that female tweets were the center of this campaign. Also the influence of the organizations on this campaign is noteworthy in a sense that one of the main factors of sustainability in an online social campaign is the high level of organization engagement to involve more people in the cause as they can mobilize other resource like events, news coverage or advertisement and increase the exposure level.

Although these results are not completely accurate, it gives a good sense of user type distribution in an online campaign in a much faster, more economical and efficient manner.

Table 3.11: Tweet, retweet and favorite distribution in entire ILLAE dataset

	Male	Female	Organization	Total
Tweet	2,995 *15.36%*	8,993 *46.14%*	7,504 *38.50%*	**19,492** *100%*
Retweet	8,464 *9.44%*	47,764 *53.28%*	33,422 *37.28%*	**89,650** *100%*
Favorite	15,830 *11.08%*	82,009 *57.41%*	45,019 *31.51%*	**142,858** *100%*

Therefore, it is a great tool for real-time monitoring of user type distribution participating during the campaign course of action and helps people specially influencers of the campaign to tune their messages.

3.6.6 Feature Importance Analysis

There are various ways to calculate and determine the importance and significance of the features on a classification task but for multi-class classification problem and determine the relative importance of each individual feature set together, I decided to train a framework (flat) with each feature set, including text, image and metadata except for name as it already gives out the individual gender(female/male). Each framework with a specific feature set except for name was trained with both datasets with imbalanced distribution to find the relative importance of each feature set to the user type classification task. The result are shown in Tables 3.12 and 3.13.

An interesting observation is that the results from both datasets are consistent. The

Table 3.12: Feature analysis results from **ILLAE** dataset with *imbalanced*(*original*) distribution

Feature	Accuracy	F_1-**Org**(%)	F_1-**Female**(%)	F_1-**Male**(%)
Name	45.08%	0%	**71.41%**	**73.20%**
Text	67.50%	82.41%	63.37%	37.14%
Image	**70.00%**	**87.77%**	57.90%	49.02%
Metadata	57.67%	70.02%	50.62%	38.85%
Proposed Framework (Flat)	**88.03%**	**95.16%**	**84.82%**	**79.07%**

Table 3.13: Feature analysis results from **CrowdFlower** dataset with *imbalanced*(*original*) distribution

Feature	Accuracy	F_1-**Org**(%)	F_1-**Female**(%)	F_1-**Male**(%)
Name	46.67%	0%	**68.66%**	**70.11%**
Text	62.02%	70.90%	65.02%	52.10%
Image	**63.42%**	**72.94%**	62.78%	57.30%
Metadata	48.70%	52.07%	56.47%	34.36%
Proposed Framework (Flat)	**75.64%**	**79.62%**	**76.65%**	**72.13%**

classifiers with image feature set outperformed other feature set classifier in terms of F_1 score of organization, 87.77% and 72.94% for ILLAE and CrowdFlower dataset respectively. It shows that organizations are more accurately identified through their profile picture than other features including name, profile description, tweets, metadata. This results prove my hypothesis that organizations tend to use non-human profile picture. On the other hand, name database is more effective and accurate in gender classification of individuals(female/male) and it outperformed all of the other feature set classifiers except the proposed framework in terms of F_1 score. This also shows that people are using their real name in their account participating in this movement.

Another interesting observation here between these two datasets is F_1 score of male class is always less than that of female class except for that of their name database where male class has a higher F_1 score than that of female. It shows that men are more easily identifiable through their profile names than women. Another point is that men do not usually follow common themes as opposed to women in terms of profile image and the content of their tweets in both datasets are not as reflective of their gender as that of women.

3.7 Limitations and Future Works

The proposed framework has some limitations, particularly in image recognition section. While investigating the performance of image feature set, the VGG-16 model, it came to my attention that the convolutional network could not learn features differentiating user type because of two important reasons. The first one was that few profile pictures of individual users were of organization type, hence the network failed to learn the discriminatory features. The second reason was that female profile pictures with short hair were classified as male. The network was unable to learn masculine features such as beard which differentiate male from female during the training.

To improve this framework, other user information like location, background profile image, the color of the user profiles and other elements of the tweets such as hashtags, URLs and media (photos and videos) could be helpful with this framework. For example, the content of the URLs and media might be relevant to the user type. Another interesting thing that also deserves the attention is to see why men are hard to be classified in Twitter in terms of the content and if there is another discriminant feature except for their name which can bring F_1 score of men closer to that of women.

3.8 Conclusion

This framework makes an important contribution to our understanding and practice of how AI and associated techniques can help those engaged in social media movements and activism campaigns. Our proposed framework make use of features from the diverse mode of information using multiple data elements in one tweet such as user profile description, name, tweet and profile images and without extra information like user's social network structure (friends/followers), historical tweets, and other additional (user mentions) information making it comprehensive and robust. Hence, the multi-modality of information provides the framework for the real-time analysis of user types and a more nuanced assessment of users participating in an online activism campaign. For leveraging social media for social good, using our framework, campaign directors can ensure that their message is going through to the right audience. This work is also applicable to other domains where identifying organizations, for instance emergency management, is important where identifying female/male participants is crucial (any campaign directed towards diversity of issues such as gender-based violence) since in these situations such as user type classification in disaster response context, time and resources are very crucial and limited, specially due to Twitter API limitation which makes it impossible to be extended to many incoming user's tweets which inevitably leads to service halt.

Chapter 4: Real-time Retweetability Classification

4.1 Introduction

In this chapter, it is attempted to propose an approach to engage targeted audience in the hashtag activism by creating more relevant and attractive content and increase the engagement level for the given activism. One of the purposes of hashtag activism is to build an online community where people can spread the news, share their experiences and exchange their ideas around the community's goal. Preece (2001) [Preece, 2001] says that there are two main components to build a successful online social community, *sociability* and *usability*. In that regard, Preece (2000) [Preece, 2000] explains that one of the three key factors that can contribute to good sociability is that the number of messages in the online community indicates how engaged people are with the community, which is in turn indicative of how well the community serves its purpose and goal. Online social engagement here is defined as sharing or reposting a message written by another person. People in online activisms like to create a message that will build traction to help sustain the movement. One of the promising characteristics of keeping people engaged in a cause is to build an online reputable social presence. However, building a social reputation requires time and planning and most of people participating in online activisms are regular people who would like to make an impact in their cause. So one of the few options available to every user is to know how to create an attention grabbing message to be able to raise awareness for the given cause. The focus of this framework is to help the activists to realize whether their messages are going to be attractive enough to engage people in the topic or not.

As discussed before in the challenges of resource mobilization, online social media, e.g. Twitter, has dynamic nature and is constantly changing which makes it extremely time consuming and laborious to be analyzed manually. Moreover the lack of real-time frameworks to help with the online activisms to spread their messages can be felt in the literature. Most of the frameworks in the literature are comprised of an offline analysis or a classification task where it needs to collect extra information to build the proposed model such as historical tweets/retweets and ego-network. Meanwhile the applicability of these models have been almost overlooked in real world situations. Due to a great deal of concern in privacy, most of social media platforms have restricted the public access to the users' information to the point that it just allows to crawl his own personal information [1] or with limitations imposed in terms of time and count thresholds to download the information [2]. These restrictions present an important challenge to the application of these frameworks. The research question here is: *"With the data collection constraints and fast paced content generation on online social media, how can we increase the social engagement of a hashtag driven activism in real time?"*

People are encouraged to participate in an online activism by posting a picture with the specific hashtag [Karbasian et al., 2018] or sharing their personal experience regarding an incident [Clark, 2016]. As mentioned earlier, creating content to be attractive enough to engage more people in the cause is one of the desired goals of activists. Text analysis as unstructured data on Twitter has been studied in various ways such topic modeling in [Naveed et al., 2011, Zhang et al., 2015b]. Moreover, one of the interesting elements

[1] `https://developers.facebook.com/docs/graph-api/overview/rate-limiting/`
[2] `https://developer.twitter.com/en/docs/basics/rate-limiting.html`

of the textual content of tweets is hashtags that make them semi-structured inside and semantically related to each other [Wang et al., 2016a, Wang et al., 2016b]. Due to the usage and creation of hashtags in hashtag activisms, I hypothesize that the cluster of some hashtags along with topics are likely to be more socially engaging and cause people to retweet.

In this chapter, I will propose a framework which only relies on the user's tweet and some historical tweets of the given hashtag activism he is going to engage by writing a tweet and predict if the tweet will be retweeted or not.

4.2 Related Works

In online social network, influencing and attracting the people to the message is one of the major attributes that have been extensively studied in the literature. These studies have been labeled differently for different research questions; retweetability [Chung, 2017], virality [Guerini et al., 2011], cascading [Cheng et al., 2014], information diffusion [Cheng et al., 2018], influence maximization [3] [Chen et al., 2009] or etc. These works approached the resharing content on online social network problem from different angels, such as the cascade size of information diffusion in the network [Cheng et al., 2014,Cheng et al., 2018, Kupavskii et al., 2012] or the influence of user's ego-network [Purohit et al., 2011, Wang et al., 2013]. These approaches were later formulated as a factor analysis or prediction task. In this section, the premise of the problem is clearly defined and can be considered as one of the retweetability problems with a defined set of constraints.

[3] It will be elaborated in this section.

4.2.1 Influence Maximization

Online social network sites, such as Facebook and Twitter, have become successful platforms to connect people and disseminate information rapidly. Their actual power lies in allowing information and ideas to influence a large population in a short period of time. However, to fully utilize these social networks as efficient information dissemination platforms for different purposes like advertising, social awareness or political change, there are multiple challenges hindering their optimal usage. One of the key challenges is finding influential individuals in a social network through whom the information can propagate faster. Imagine the following scenario for this problem. A social activist comes up with an idea to address a social issue that hasn't been raised recently. She would like to create a momentum in an online world, e.g. online social network, where she could attract people's attention to the cause. But she has limited resources like money or connections to influential users. She comes up with an incentive (*e.g. financial reward or online social promotion*) to be able attract a limited number of people. She hopes that these people would take this opportunity and start promoting the given cause in his/her circle of friends leading to the message propagation in the network. Now the problem becomes who to choose to influence largest number of people in the network referred to as influence maximization. With the help of these online social media platforms, the problem can be mitigated easily but it has also its own challenges. The social networks are large-scale, have complex connection structures, and have dynamic nature which means that the solution to the problem needs to be very efficient and scalable.

Domingos and Richardson [Domingos and Richardson, 2001, Richardson and Domingos, 2002] are the first to study influence maximization as an algorithmic problem. Their methods are probabilistic, however. Kempe, Kleinberg, and Tardos [Kempe et al., 2003] are the first to formulate the problem as the following discrete optimization problem. A social network is modeled as a graph with vertices representing individuals and edges representing connections or relationship between two individuals. Influence are propagated in the network according to a stochastic cascade model. Three cascade models, namely the independent cascade model, the weight cascade model, and the linear threshold model, are considered in [Kempe et al., 2003]. Given a social network graph, a specific influence cascade model and a small number k, the influence maximization problem is to find k vertices in the graph (refered to as seeds) such that under the influence cascade model, the expected number of vertices influenced by the k seeds is the largest possible. Kempe et al. presents a greedy approximation algorithm applicable to all three models. They also show through experiments that their greedy algorithm significantly outperforms the classic degree and centrality-based heuristics in influence spread.

Chen et al .[Chen et al., 2009] tackled the efficiency issue of influence maximization by improving the algorithm proposed by [Kempe et al., 2003] and also by proposing a more efficient and faster heuristic for determining the influence spread.

The problem formulation of influence maximization present two issues in the context of efficient resource mobilization for online social movement. Firstly due to the limitation imposed by these online social media, it is impossible to be able to construct a social network graph whose individual's influence spread will be evaluated. Secondly due to the limited resources available to individuals or activists willing to promote their cause, financial benefit or similar rewards will be unavailable to be utilized. Those being said, the influence maximization approach cannot address the issue explained earlier for such movements.

4.2.2 Retweetability

In the literature, it can be argued that the user's online social status is more important that the message itself [Cheng et al., 2014] or the content of the message is an important component to have a positive effect in spreading in the online community as well [Petrovic et al., 2011]. Some researchers focused on the types or content of tweets that are more likely to be retweeted than others [Rudat and Buder, 2015, Rudat et al., 2014, Son et al., 2013, Sutton et al., 2015a, Sutton et al., 2015b, Chung, 2017, Zaman et al., 2010]. Research also has shown that the more frequently retweeted tweets are the ones that tend to be sentimental and emotional, such as tweets that include news, describe negative consequences or emotions or remind personal experiences [Rudat and Buder, 2015, Rudat et al., 2014, Stieglitz and Dang-Xuan, 2013].

Event-centric hashtag analyses of tweets such as disastrous events, e.g. hurricane and terrorist attack, show that tweets describing impacts from hazards were frequently retweeted, whereas tweets expressing gratitude were less retweeted [Sutton et al., 2015a, Sutton et al., 2015b] and also information including phone number, incident-related data, date or time regarding an update make the tweets in these events more important to be retweeted [Neppalli et al., 2016].

Other researchers studied technical factors (such as external links, hashtags, images, videos, and the use of mentions in tweets) to find their effects on retweetability. Their findings showed that tweets with hashtags [Bao et al., 2013, Son et al., 2013, Suh et al., 2010, Sutton et al., 2015a] and multimedia, such as pictures, videos, and emoticons, were more likely to be retweeted [Huang and Xue, 2015]. There were some studies around how retweeting behavior was influenced by user- and network-related factors [Purohit et al., 2011]. They showed that a message was more likely to be retweeted when the message is in the follower's areas of interests [Luo et al., 2013,Yang et al., 2010,Zhang et al., 2016,Wang et al., 2013] and when the message was retweeted frequently by many followers [Bao et al., 2013, Zhang et al., 2015a]. In the unstructured part of tweet content, text understanding has been studied extensively but using latent features like topic helped with the findings of interesting topics in tweets [Naveed et al., 2011, Zhang et al., 2015b].

In Tables 4.1 and 4.2, you will see the list of *real-time* and *non real-time* features used and proposed in the 12 related papers that were discussed earlier along with the features of our proposed framework (column F) which will be discussed in the following sections. They are numbered from 1 through 12: 1-[Suh et al., 2010], 2-[Petrovic et al., 2011], 3-[Naveed et al., 2011], 4- [Purohit et al., 2011], 5- [Luo et al., 2013], 6- [Wang et al., 2013], 7- [Zhang et al., 2015a], 8- [Pervin et al., 2014], 9- [Zhang et al., 2015b], 10- [Zhang et al., 2016], 11- [Neppalli et al., 2016] and 12- [Chung, 2017].

In these tables, the features are categorized primarily in terms of real-time and non real-time accessibility. Then they are broken down to the types of features into user, content, network and temporal groups. There are three types of feature values in the table: binary(B), categorical (C) and numerical (N). For the description of each feature, please refer to the respective paper.

4.3 Contributions

The main contribution in this framework is a method to determine in real-time based on the history of a given hashtag activism, in this case, two engineering diversity movements, the *ILookLikeAnEngineer* and *WomenInEngineering* movements on Twitter, if a tweet that a user is writing in the given hashtag activism will be retweeted/reshared. It provides a tool for relevant content creation of a tweet in a hashtag campaigns leading to raising more

Table 4.1: List of **Real-Time** Features in Related Works and Our Proposed Framework

	Type	Feature	1	2	3	4	5	6	7	8	9	10	11	12	F
Real-Time	User	Status	N	N			N						N		N
		Favorites	N	N									N		N
		Age of the Account	N						N				N		N
		Followers	N	N			N		N	N			N	N	N
		Followees	N	N			N		N				N	N	N
		Verified		B			B		B				B		B
		Listed Count		N									N		N
		English		B											
	Content	Hashtag Presence		B	B					B			B		B
		Hashtag Count	N			N							N		N
		Average Hashtags Length													N
		Mention	N	B	B	N	N								B
		URL	N	B	B	N				B			B	B	B
		Media													B
		Photo												B	
		Video												B	
		Word Count												N	N
		Character Count		N									N		N
		Reply		B	B								B		B
		Sentiment (LIWC)-93 features				N									N
		One Word Sentence											B		
		Phone Number											B		
		Measuring Units											B		
		Date or Time											B		
		Emoticons											B		
		Cusswords											B		
		Keywords-Manually Selected				N							B		
		Abbreviations											B		
		Trending Words		C											
		Novelty Score		N											
		Bag-Of-Words		C											
		Creation Time		C											
		Temporal Weight				N									
		Topics (LDA)			N						N				N
		Exclamation Mark			B										
		Question Mark			B										
		Term Positive			B										
		Term Negative			B										
		Emoticons Positive			B										
		Emoticons Negative			B										
		Sentiment (ANEW)-3 features			N										
		Terms			N										
		Tweet Embedding										N			
		Clusters of Hashtags (LDA)													N
		Communities of Hashtags													N

Table 4.2: List of **Non Real-Time** Features in Related Works and Our Proposed Framework

	Type	Feature	1	2	3	4	5	6	7	8	9	10	11	12	F
Non Real-Time	User	Type												C	
		Active Followers				N			N						
		Active Followees				N			N						
		Klout Score				N									
		Gender							B						
	Content	SimInterest					N								
		Previous Retweet Count from a given user					N								
		Previous Mention Count from a given user					N								
		Manual Coding												C	
		Retweet Count									N				
		Retweet History						N			N				
		User's Tweets Embedding										N			
		Author's Tweets Embedding										N			
		Topics of Historical Tweets (LDA)							N						
	Network	AmplifierScore								N					
		InformationStarterScore								N					
		TransmitterScore								N					
		PeopleAware(t-1)								N					
		AveragePosition								N					
		Weakly Connected Component				N									
		Community Size				N									
		ego Network						N	N						
	Temporal	TimeOfDay								C					
		DayOfWeek								C					
		Age								N					
		Same TimeZone					B								
		PostTimeConsis					N								

awareness in the cause.

The mutual relationship between topic and hashtag was studied in sub-event discovery and found promising [Xing et al., 2016]. I used this relationship to my advantage to improve my retweetability framework which only relies on just one tweet information and the user's meta data and utilizes the latent features in the textual content of the tweet. With

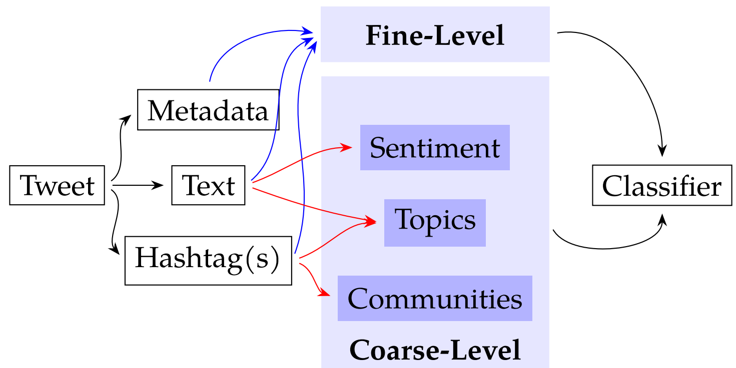

Figure 4.1: Proposed framework for retweetability classification with fine-level and coarse-level content-based features

the analysis of hashtags used in the tweets, our findings suggest that users using a specific set of hashtags and topics are more likely to attract more people in the cause.

And also the proposed framework only relies on one tweet and requires no further structural or historical information about the network and users. This characteristic makes it interestingly applicable in real-world situations where it delivers the results in real-time by ignoring the aforementioned limitations in social media platforms.

4.4 Proposed Framework

Here, I propose a framework that uses information from a user's one tweet and user's profile information to classify the tweet as retweetable or not. This framework, Figure 4.1, uses two types of features: fine-grained- and coarse-grained- content-based features [Naveed et al., 2011]. In this section, every component of the framework will be explained along with the proposed framework whose feature sets can be found in Tables 4.1, 4.2 under column *F*.

4.4.1 Fine-Grained Level Features

Fine-grained level features, or fine-level features, are constructed by extracting directly from the tweet JSON and do require little or no processing [Naveed et al., 2011]. In tweet JSON, there is a set of metadata stored that could help with the task of classification, but some of them have been found quite helpful in the literature. Those data related to user have been studied extensively and proven to be relevant to the task including statuses (tweets) count, favorites count, followers count, followees (friend) count, listed count, age of the account and verified status which is the most influential factor in this particular task. If a user is listed many times, i.e., many lists follow him, this should mean that he tweets about things that are interesting to a larger user population and if the account is verified, it shows that that almost anything that celebrities write will get retweeted, and thus having this feature should improve performance [Petrovic et al., 2011]. From the text, word count, URL and media presence are included. And it also includes if the tweet has any user mention and if it is a reply to another tweet.

From the hashtag(s), three features are extracted: if the tweet has any hashtag, how many hashtags it has and the average length of hashtags. These features are easily calculated from any given tweet and will be given as a part of features for the classification task. Eltantawy et al. (2011) [Eltantawy and Wiest, 2011] showed that using other hashtags helps the movement grow as it is connecting to another similar movement like Egyptian and Tunisian revolution. They can share the lessons learned through this exchange. This is the similar case that we see here in *WomenInEngineering* and other similar hashtag driven activisms like *WomenInSTEM* and *WomenInScience*. It reinforces the idea that the set of particular hashtags being used can increase the probablity of getting retweeted in the network.

4.4.2 Coarse-Grained Level Features

Contrary to fine-grained level features, coarse-grained level features, or coarse-level features, are constructed by the help of external information, e.g. human (LIWC) or pre-trained model [Karbasian et al., 2018], or another source of data, e.g. topics [Naveed et al., 2011] or tweets embedding [Zhang et al., 2016]. These features create a higher level of abstraction from the data that are more interpretable and understandable by human (e.g. topics or emotional sentiment). In the following, each coarse-level feature generation methodology will be explained.

Linguistic Inquiry and Word Count (LIWC)

Creating sympathy in the messages can be useful to attract more like-minded audience to join the movement. It is associated to increase the social reward as it captures emotional and psychological themes of the messages and encourage people to chime in and feel included in the cause. Studies showed that emotional engagement are effective persuasive devices. In the context of written communication, previous research has indicated that emotional stimuli in terms of emotion words or emotional framing of messages may increase the level of attention [Bayer et al., 2012, Kissler et al., 2007, Smith and Petty, 1996]. Hence increased attention may in lead to a higher likelihood of behavioral response to emotional stimuli in terms information sharing [Heath, 1996, Luminet IV et al., 2000, Peters et al., 2009, Rimé, 2009]. Emotional content suchs positive (awe) or negative (anger or anxiety), emotions have been found more viral. It can be argued that increased level of attention triggered by by emotions in written communication are determinants of sharing behavior.

As explained in the previous chapter, LIWC2015 was used for the textual part of the tweet to capture the features specially sentiment and psychological aspects of the tweet to help with the retweetability prediction task.

Topic Modeling

Topic modeling has been used to find topics (cluster of words) in texts before and allowed us to gain a better insight into what the texts are talking about at a higher level. In this framework, I used the popular topic modeling technique called latent Dirichlet allocation (LDA) [Blei et al., 2003], as it has been proven effective for finding discussion topics in natural language text documents [Naveed et al., 2011]. LDA works as a clustering algorithm to find semantically relevant words in the corpus and put them in a cluster of words (topic). It is an open vocabulary approach to text analysis as it only relies on the documents at hand and find topics according to them as opposed to the closed vocabulary approach, LIWC, where predefined word-category lexicon is used to determine the sentiment and psychological tone of the document. Using these two approach together have been found quite helpful to understand the text [Schwartz et al., 2013b]. Also as tweets are usually short and they suffer from sparsity issue in topic modeling methodologies, hashtags could serve as weakly-supervised information for topic modeling, and the relation between hashtags could reveal latent semantic relation between words [Wang et al., 2016b]. Therefore, I hypothesize that using the cluster of various co-occurring hashtags shared in a social activism campaign helps people from other similar movements to connect where they can share the lessons learned through this exchange. The idea here is to apply topic modeling to hashtags and find relevant clusters of hashtags that could contribute to retweetability.

Given a set of documents, LDA uses machine learning algorithms to infer the topics and topic memberships for each document. In this framework, I used the implementation of the LDA model provided by MALLET version 2.0.8 [McCallum, 2002] [4], which is an implementation of the Gibbs sampling algorithm [Geman and Geman, 1987]. The number of cluster like other clustering algorithms should be prespecified as it falls into the hyper-parameter tuning part of the framework that will be discussed later. The result of LDA is (a) a set of topics, defined as distributions over the unique words in the tweets and (b) a set of topic membership vectors, one for each tweet, indicating the percentage of words in the tweet that came from each topic. As mentioned before, the highest-probable words in a topic are semantically related, which together reveal the nature, or concept, of the topic. This vector shows that LDA is a soft clustering algorithm that each tweet can have probability of each topic.

LDA discovers K topics, $z_1, ..., z_k$. I denote the membership of a particular topic z_k in a document, i.e. the text of the tweet or the list of hashtags here, d_i as $\theta(d_i, z_k)$. I note that $\forall i, k : 0 \leq \theta(d_i, z_k) \leq 1$ and $\forall i : \sum_k^1 \theta(d_i, z_k) = 1$. Then I define a threshold, σ, to indicate whether a particular topics is "in" a tweet. Usually, a document will have between 1 and 5 dominant topics, each with memberships of 0.10 or higher [Blei et al., 2003]. However, due to the probabilistic nature of LDA, sometimes topics are assigned small but non-zero, e.g., 0.01, memberships to a document, and are can be considered noisy. Thus, by using the σ threshold as a membership cutoff, I keep only the main topics in each document and discard the probabilistic errors. Then for each document, I normalize the weight of topics to be 1. σ is a hyper-parameter of LDA that will be specified later.

[4] `http://mallet.cs.umass.edu`

Community Detection

Community detection is a clustering algorithm for graph-based problems. As opposed to topic modeling, community detection is a hard clustering algorithm where each data point either belongs to a cluster or community completely or not. Each cluster is a set of nodes in the graph that have more edges linking among its members than edges linking outside to the rest of the graph [Gargi et al., 2011]. To construct the graph with text, the co-occurrence matrix of words can be used. The value of each item in that matrix can be either binary or weighted. Due to the sparsity issue of text-generated graph as discussed earlier, I used hashtag-generated graph then applied Louvain community detection algorithm[5] to find the communities of the hashtags that will be used for retweetability classification problem. The feature extracted from communities will be the vector whose each item represents each community. The value of each item is the presence of that community in the tweet; either binary or weighted. The number of clusters In Louvain algorithm is determined by the algorithm using maximizing each community's modularity metric ($Best\_Louvain$).

4.4.3 Model Fitting

As there are two possible outcomes of retweetability and all of the aforementioned features take either real or binary value, the retweetability problem is treated as a binary classification problem operated on feature vectors of the following format.

- **Label**: The fact of whether the tweet will be retweeted by another person or not. The value is a binary variable which can be either positive or negative, and it serves as the class label.

[5] **Louvain** algorithm is selected because of efficiency for large networks.

- **Fine-grained Level Features**:
  - *Status*: Number of user's tweets
  - *Favorites*: Number of user's favorites
  - *Age of the Account*: Number of days after the creation of user's account
  - *Followers*: Number of user's followers
  - *Followees*: Number of user's friends (followees)
  - *Verified*: If the user is verified or not
  - *Listed Count*: Number of lists that user is being followed by
  - *Length*: Number of characters in the tweet
  - *Word Count*: Number of words in the tweet
  - *Hashtag Count*: Number of hashtags in the tweet
  - *Average Hashtags Length*: Average length (number of character) of hashtags in the tweet
  - *URL*: If the tweet has url or not
  - *Media*: If the tweet has photo or video or animated gif or not
  - *Mention*: If the tweet has any user's mention or not
  - *Reply*: If the tweet is a reply to another user's tweet or not
- **Coarse-grained Level Features**:
  - *LIWC*: 93 numerical values of the tweet from LIWC
  - *Topics (LDA)*: K [6] numerical values representing topic membership weights for the text of the tweet from LDA

[6] It will be assigned in the hyper-parameter tuning section.

– *Clusters of Hashtags* (*LDA*): K [6] numerical values representing cluster weights for the list of hashtags in the tweet from LDA

– *Communities of Hashtags*: L[7] numerical values representing cluster weights for the list of hashtags in the tweet from Louvain algorithm

4.5 Datasets

For retweetability classification task, I used two datasets; both of which are hashtag activsms for women's issues in male dominant engineering world. The brief description is provided for each in the following sections. You can see the statistics of both datasets in Table 4.3.

4.5.1 ILookLikeAnEngineer (ILLAE)

The first dataset about hashtag activism toward engineering diversity is ILookLikeAnEngineer campaign, as explained in the introduction section.

4.5.2 WomenInEngineering (WIE)

There are many gender equity and women empowerment movements created by women in STEM field such as WomenInEngineering, WomenInSTEM, WomenWhoCode. #WomenInEngineering was first used by an American student in April 2010 but it never caught on until June 2013 when there was an article [8] about Roma Agrawal, structural Engineer and STEM advocate, who was trying to change the perceptions of the industry.

[7] It is determined by the algorithm, $Best\_Louvain$ number of communities maximizing graph's modularity metric.

[8] `https://www.theguardian.com/artanddesign/architecture-design-blog/2013/jun/26/women-in-engineering`

Table 4.3: Statistics of *ILookLikeAnEngineer* and *WomenInEnginnering* datasets

		ILLAE	**WIE**
	Time Span	Aug 8,2015-Oct 15,2015 (∼2 months)	Apr 29, 2018-Sep 9,2019 (∼16 months)
All	**Tweets**	17354	8225
	Users	11828	3,748
	Tweets/User	1.47	2.19
	Retweet Count	26,398	13,020
	Retweeted Tweets	5,168 (29.77%)	4,288 (52.13%)
Verified	**Tweets**	1,114 (6.42%)	565 (6.87%)
	Users	457 (3.86%)	167 (4.45%)
	Tweets/User	2.44	3.38
	Retweet Count	9,169 (34.73%)	1,949 (14.96%)
	Retweeted Tweets (verified-total)	853 (76.57%-16.5%)	411 (72.74%-9.58%)

I collected the dataset based on #WomenInEngineering hashtag for almost 16 months recently and unlike ILLAE, it is during the movement when the initial surge has been relaxed.

4.5.3 Exploratory Analysis

In both movements, verified users were more active than non-verified users in terms of writing tweets and getting retweeted. Although the social reputation, i.e. verified status in Twitter, increases the probability of retweetability in both cases, more than 72% as opposed to the overall retweeted percentage, the majority of the users, more than 94% users, are non-verified and that reinforces the idea that the framework for non-verified or regular users is needed to help them know how engaging their tweets should be to be retweeted.

As you can see in Table 4.3, verified users were more active than the regular users in ILLAE.

As the task is a binary classification problem, to avoid the imbalance problem in training, we made both datasets balanced by random sampling from the majority class as many as the other class. In this case, for ILLAE dataset the number of samples per each class is 5168 and for WIE dataset is 3937.

4.6 Experiments

4.6.1 Preprocessing

Cleaning the textual data is the most important preprocessing part of text analysis. Hence, I cleansed the textual content of the tweet in multiple steps. First, I removed any HTML tags, URLs, emails, user mentions and hashtags. Then I applied the Porter stemming algorithm [Jones, 1997]. The next step was to remove common English-language stop words. Since 2-grams (equivalently, bi-grams) have been shown to increase the quality of text analysis [Tan et al., 2002], I included the bi-grams as well for LDA model. To get more meaningful words for each tweet, I used lemmatization technique which identifies the intended *part of speech* and meaning of a word in a sentence. So I just used adjective, adverb, noun and verb as accepted parts of speech. The final step was to remove less and highly frequently used words from the tweets. I set the minimum threshold (*no_below*) for each word and a percentage (*no_above*) of the documents for the maximum threshold. Once all features are generated as described before, the values of features are standardized using z-score. You can find list of hyper-parameters for each feature in Table 4.4 at *Feature* section.

Table 4.4: List of Hyper-Parameters Used in the Experiments

	Type	Hyper-Parameter
Feature	**LDA**	k = [25, 50, 75, 100] no_below = [1, 10] no_above = [0.5, 1] σ = [0, 0.1]
	Communities	k = [25, 50, 75, 100] edge type = [binary, weighted] value = [binary, weighted]
Algorithm	**SVM**	kernel=[linear, poly, rbf, sigmoid] cost = [0.01, 0.01, 0.1, 1, 10, 100, 100] gamma = [0.001, 0.01, 0.1, 1] coef0 = [0.1, 1, 10, 100] degree = [2, 3, 4]

4.6.2 Support Vector Machines as Classifier

The features after preprocessing become ready to train the SVM classifier for the prediction task like the previous chapter. The hyper-parameters for SVM used in the experiments are presented in Table 4.4 at *Algorithm* section.

4.6.3 Model Selection

Hold-out cross validation was used for training and testing the framework and the baselines. 20% of the dataset was held out for the final testing and the rest of it (80%) was used for 5 fold cross validation. For performance measure in hyper-parameter optimization for each CV phase, "accuracy" was used. Then based on the best set of hyper-parameters in cross validation phase, the training set was used for training then it was tested against the held out test set. "Accuarcy" and "F1 score" for positive class were used for performance measure of each model. To compare each classifier, I used McNemar's test [McNemar, 1947] and the p-value threshold was set 0.05.

4.6.4 Baselines and Frameworks

To see the relative importance of each feature set, the following is the set of features from proposed framework:

- *All*: contains all fine-level features and coarse-level features.
- *Content-All*: contains just all of the content features.
- *Content-Coarse*: contains just coarse-level content features.
- *Content-Hashtags*: contains just clusters of hashtags and communities of hashtags.
- *Content-Topics*: contains just topics (LDA).
- *Content-LIWC*: contains just content features from LIWC.
- *Content-Fine*: contains just fine-level content features.
- *User*: contains just user's metadata features.

The baselines for this prediction task that have only real-time features from Tables 4.1, 4.2 are as the following:

- *Baseline-1*: column 1 [Suh et al., 2010]

- *Baseline-2*: column 2 [Petrovic et al., 2011]

- *Baseline-3*: column 3 [Luo et al., 2013]

- *Baseline-4*: column 11 [Neppalli et al., 2016]

4.6.5 Experimental Results

The results of *ILookLikeAnEngineer* and *WomenInEngineering* are demonstrated in Figure 4.2 and 4.3 respectively. They show the accuracy and F1 score for positive class achieved by SVMs on 4 baselines and feature sets based on the best hyper-parameters in the model selection phase.

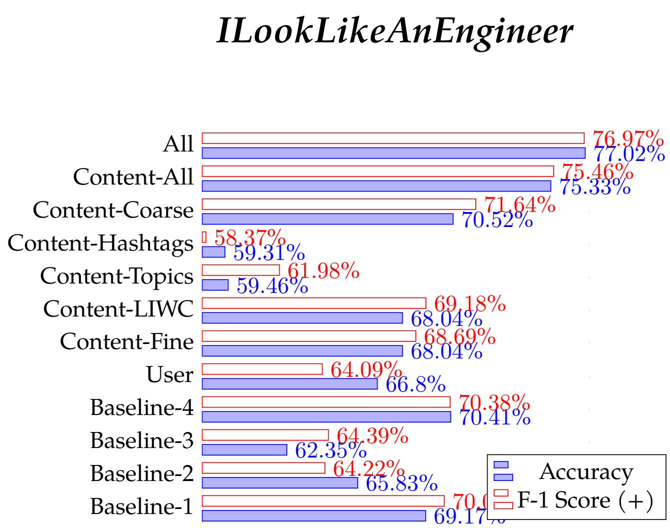

Figure 4.2: Experimental results of 4 baselines with different sets of features from proposed framework for *ILookLikeAnEngineer* dataset

The statistical comparison results with McNemar test were that the *framework* (*All*) against 4 baselines in both datasets were statistically significant ($p < 0.05$).

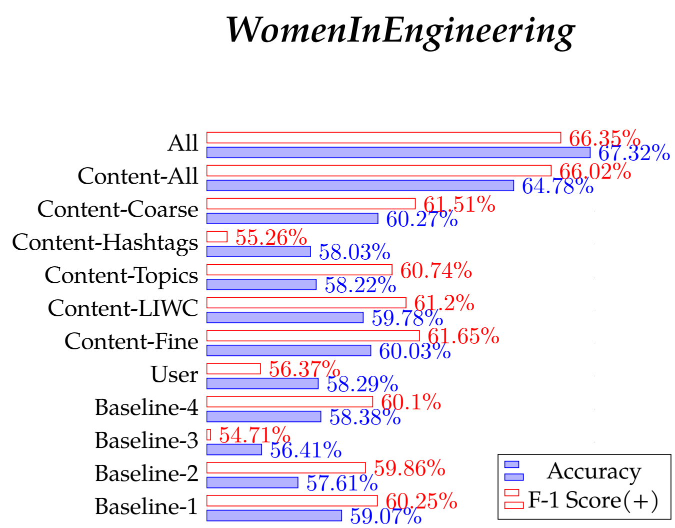

Figure 4.3: Experimental results of 4 baselines with different sets of features from proposed framework for *WomenInEngineering* dataset

Our observations on experiments are listed as following:

- I observe that the performance of the proposed framework (*All*) outperforms all of baselines with statistically significant difference and increase of accuracy between *framework*(*All*) and the best baseline were 6.61% on ILLAE dataset and 8.25% on WIE. It shows that adding relevant features to the task can contribute to the the performance improvement. In this case, due to the nature of hashtag driven activism, the abundance of hashtags could provide a set of latent features in terms of their mutual relationship that were created through coarse-level feature extraction methodologies described earlier.

- Contrary to the findings in [Petrovic et al., 2011], just content-related features (*Content-All*) are more indicative of retweetablity potential than just user-related features

(*User*) which is online social status and the social network connections. It shows the fact that the messaging and content in these two hashtag activisms are major contributing factors that need to be looked into in more details.

- I also observe that there is no statistical difference between coarse-level features including sentiment (*Content-LIWC*), topics (*Content-Topics*) and hashtags (*Content-Hashtags*) in both datasets except for ILLAE where hashtags features are weaker than user features but comparable. It is interesting that each coarse-level feature was able to capture a latent aspect of tweets content and alone they are comparable to user's social network status. It seems like more latent features in this direction will be more helpful with this task as well.

- Comparing the fine-level content features (*Content-Fine*) with coarse-level content features (*Content-Coarse*), there is also no statistical significance. As explained earlier, fine-level features are the metadata given in tweet JSON and there is little computation done to extract them. On the other hand, coarse-level features are a higher abstraction of the the text of the tweet. Understanding more coarse-level features, i.e. latent features in different modality of the data, such photo, video and hyperlinks is definitely more helpful in such task.

4.6.6 Feature Importance Analysis

There are various ways to calculate and determine the importance and significance of the features on a classification task. The standard approach to answer this question is to use regression analysis for a binary classification problem. Since the retweetability problem here is one of them, the suggested version of regression analysis is to use logistic regression. [Naveed et al., 2011]. Logistic regression is a generalized linear regression method for learning a mapping from any number of numeric variables to a binary or probabilistic variable [Hosmer Jr et al., 2013]. In equation below, you see the formula of logistic regression:

$$P(tweet_j\ is\ retweetable \mid f) = \frac{1}{1 + e^{-(w_0 + \Sigma_i w_i f_{ij})}} \tag{4.1}$$

f is the feature vector of the framework and the weights, w_i, are associated to each feature. As all of the features for the classification task have been z-score normalized, the weights can be compared to each other. If a weight is around zero, the respective feature has almost no effect on the retweetability whereas any positive and negative weight have positive and negative effect on the task respectively. The results of the regression analysis are presented in from Tables 4.5 through 4.10. They are broken down in the same group as the features were explained earlier; fine-grained and coarse-grained.

Table 4.5 shows the weights of fine-level features. As LIWC has 93 features, I just show the weights of the first and last three features in Table 4.6. Tables 4.7 and 4.8 are the weights of the first and last topic along with the top keywords and one sample tweet for both datasets; *ILLAE* and *WIE*. And finally Tables 4.9 and 4.10 are the weights of first and last hashtag cluster using LDA with top hashtags and one sample tweet for both datasets as well. In the following, I present my observations from the regression analysis:

- In Table 4.5, *status count* is one of the best indicators of retweetablity in both datasets saying that the more status count, the less likely your tweet will be retweeted.

- In *ILLAE*, *Followers Count* has the largest weight showing that the more follower you have, the more likely your tweet will get retweeted. Although *Followers Count* has positive weight in *WIE*, but *Length* has the largest value. On another note, *Listed*

Table 4.5: List of Weights for Fine-Level Features Learned from Logistic Regression for *ILookLikeAnEngineer* and *WomenInEngineering*

Feature	ILLAE	WIE
Age of Account	-0.132	-0.043
Avg Hashtags Length	0.014	0.06
Favorites Count	0.174	0.228
Followers Count	**3.358**	0.147
Friends Count	0.076	0.088
Hashtags Count	-0.001	0.067
Length	0.369	**0.387**
Listed Count	1.459	0.035
Media	0.494	0.252
Mention	0.254	-0.181
Reply	-0.159	0.036
Status Count	*-1.008*	*-0.1*
URL	-0.16	0.245
Verified	0.351	-0.193
Word Count	-0.07	-0.428

Count is the second largest weight in *ILLAE* and *Media* is the second largest weight in *WIE*. I can deduce that *ILLAE* movement was more driven by the more famous and well-known people who merely used their social media reputation to promote the movement message, [Johri et al., 2018b], whereas in *WIE*, the content has more importance than user's social media network.

- Table 4.6 shows the features from LIWC. To further study each feature, please refer to their manual[9]. An interesting observation between these datasets is that the favorable tone of language being used in the tweets for retweetablity is *informal* whereas *netspeak*, e.g. btw, lol, thx, has negative effect.

[9] `https://liwc.wpengine.com/wp-content/uploads/2015/11/LIWC2015_LanguageManual.pdf`

Table 4.6: List of Weights for the First and Last 3 Features in LIWC for *ILookLikeAnEngineer* and *WomenInEngineering*

ILLAE		**WIE**	
Feature	**Weight**	**Feature**	**Weight**
function	0.559	informal	0.67
pronoun	0.307	affect	0.345
informal	0.26	female	0.196
⋮	⋮	⋮	⋮
netspeak	-0.306	assent	-0.285
affect	-0.372	posemo	-0.481
cogproc	-0.45	netspeak	-0.513

Table 4.7: List of Weights for the First and Last Features in Topics for *ILookLikeAnEngineer*

Topic	**Weight**	**Keywords**
Challenge News Ad	0.064	tech, challeng, women, stereotyp, industri, news, big, race, brave
Tweet: #Ilooklikeanengineer Is A Big, Brave Challenge To Tech Industry Stereotypes `http://t.co/Vbcskyrs5R` #News #Race #Women #Socialmedia		
⋮		
Spam News Ad	-0.067	divers, campaign, ad, aim, tech, challeng, promot, preconcept, spark
Tweet: Promoting diversity by challenging preconceptions: The #ILookLikeAnEngineer ad campaign aims to dispel ... `http://t.co/6oo2s2Q5nE`		

- Tables 4.7 and 4.8 are for the topics from LDA. The topic for the *Spam News Ad* was negatively correlated to the retweetability as there were many tweets with the same content and they were considered spam by the users, even though the topic and the keywords are relevant.

- Another attempt by a company using the hashtags for the movement can be seen in

Table 4.8: List of Weights for the First and Last Topics for *WomenInEngineering*

Topic	Weight	Keywords
Robot Research	0.098	womeninengin, microsit, read, research, womenintech, rail, find, articl, robot, contribut
Tweet: Vision-based control is a technique that uses visual features extracted from images to control the motion of a robot, such as #drone navigation - find out more on this topic via our #WomenInEngineering microsite: `https://t.co/kQno5DQuYJ` `https://t.co/mXmozPKHgP`		
⋮		
Doll Company	-0.072	work, engin, teweek, lottietour, lotti, lab, week, meet
Tweet: Lottie trying out a dream job at Aston Martin! #LottieTour #WomeninEngineering #WES `https://t.co/wt8frkggZg`		

Table 4.8 in the last feature. A doll company, called *Lotti*, used the hashtag to promote its products but its tweets in the movement didn't attract people a lot.

- Tables 4.9 and 4.10 are the weights for the hashtags clusters from LDA. *Challenge News Ad* in Table 4.10 has negative weight in terms of hashtags cluster as opposed to the Table 4.7 where it has the positive weight in terms of textual content. It shows that even though the relevant keywords were used to promote the site and products, the tweets were less likely to get retweeted whereas STEM related cluster of hashtags like **INWED18**, *International Women In Engineering Day 2018*, was more likely to get retweeted.

4.7 Limitations and Future Works

Using the cluster of hashtags as one of its feature sets, this framework has the limitation of being extended in other online form of activism except for hashtag-driven ones.

This work can be improved by the following points:

Table 4.9: List of Weights for the First and Last Hashtags Clusters (LDA) for *ILookLikeAnEngineer*

Topic	Weight	Hashtags
Employee Intro	0.119	inwed18, raisingthebar, womeninstem, engin, stem, transformthefutur, yoe, divers, iwe
Tweet: Meet Francesca from #GE #OilandGas #Florence #Italy #ILookLikeAnEngineer #LooksLikeAnEngineer http://t.co/qGqiIX5wbk		
⋮		
Challenge News Ad	-0.146	women, socialmedia, news, race, addwomen, twitter, winconf, nuclear, digit, jasperengin
Tweet: #Ilooklikeanengineer Is A Big, Brave Challenge To Tech Industry Stereotypes http://t.co/Vbcskyrs5R #News #Race #Women #Socialmedia		

Table 4.10: List of Weights for the First and Last Hashtags Clusters (LDA) for *WomenInEngineering*

Topic	Weight	Hashtags
INWED18	0.102	inwed18, raisingthebar, womeninstem, engin, stem, transformthefutur, yoe, divers, iwe
Tweet: Two weeks to go ! #WomenInSTEM #RasingTheBar #WomenInEngineering #SheCanEngineer #INWED2018 #YearOfEngineering #STEM #ThisIsEngineering #INWED INWED1919 https://t.co/rWZLIEq8Iv		
⋮		
Job Ad	-0.072	maceng, macengphi, engineeringphys, physic, thinkengin, bigidea, flexiblework, womeninstem, glasgow
Tweet: NEW JOB 2 Project Civil Engineers needed to join Harleyhaddow in #Edinburgh ; #Glasgow! The roles are 4 days/wk with plenty of flexibility. APPLY NOW or please share! https://t.co/ns0dSXGkWh #flexibleworking #womeninengineering #womeninSTEM #Glagsowjo... https://t.co/d1jYUH2EZF		

- As seen in the experimental results, fine-level content features are a great potential to explore as they have been studied before in different tasks [Li et al., 2016]. Understanding media, photo and video along with the tweet as well as the analysis of hyperlinks can be included as a set of coarse-level content features to increase the performance specially where using of photo in the tweet is encouraged in the movement [Johri et al., 2018b].

- Petrovic et al. (2011) [Petrovic et al., 2011] shows that the hour of tweeting affects the retweetability probability. In this regard, the temporal aspect of the tweet can be considered as an element of improvement to the task.

- Although this framework is focused on the applicability in the online social media platforms' current policies, the next step can be a recommender system which recommends highly probable tweets to a user based on his social network metadata and his recent tweets in the given activism. This system helps users to come up with a retweetable content which will eventually help the given movement grow in the network and spread the messages.

- SVM was used as the classification algorithm due to it non-linearity capability. But a specific classification framework can be devised using all of the relevant features with the help of deep neural network such as the work in [Zhang et al., 2016] as it has been proven useful to different variety of classification tasks.

4.8 Conclusion

In this chapter, I presented a real-time retweetability prediction framework to help a hashtag-driven activism spread the messages in the social network. This framework was tested against two hashtag activism datasets and outperformed all of 4 baselines with more than 6.61% and 8.25% accuracy increase. This framework only relies on the given tweet data to make the prediction and is deemed practical and real-time. This framework instead of getting extra information about the user's social network and his historical tweets that cannot be accomplished by the limitation imposed by online social platforms (in this case, Twitter), focuses on the features available in the content of the tweet and tries to extract latent features helping the prediction task. Focused on hashtag-driven activisms, it tries to find a set of features relevant to the data and in this case hashtags which eventually helps with the retweetability task. This framework helps messages propagate in the given hashtag activism more efficiently and faster.

Chapter 5: Discussion, Future Works and Applications

5.1 Discussion

The focus of this dissertation was to study the issues inhibiting the growth of engineering diversity related hashtag activisms and propose solutions to improve the resource mobilization in such movements. This work provided two analytical frameworks to address the challenges facing hashtag campaigns in terms of user participation and messaging. Those frameworks were intended to help the activists and regular users to raise awareness through online social movement for their desired audience such as families, students or female engineers to create a support system to keep them motivated in these fields. Moreover these two real-time frameworks can be extended and utilized in other hashtag campaigns as long as the underlying hypothesis for each users are the same. These frameworks improve the ability of a diverse range of people to participate in online campaigns and to have their opinions heard in a more economic, efficient and faster approach. The core of this work is in a direction toward a broad support for social good.

Each preceding chapter addressed the following goals:

1. It first studied STEM and identified the lack of URM diversity as a major issue, then reviewed the existing solutions. Later social movement and its principal components were studied. Then hashtag activism campaign was examined as a facilitator of the change for diversity issue. Next the challenges for hashtag activism campaigns were

Table 5.1: List of challenges and solutions for hashtag activism campaigns

	Challenges	**Solutions**
Resource Mobilization Online Social Media	• Dynamic nature and constantly changing. • Information overload.	• Machine Learning Algorithm. • Real-Time Analysis: no extra information required.
Activities Writing and sharing a tweet	• Clicktivism, slacktivism or engaged passivity • Lack of collective identity and solidarity	• Use of hashtag for SM (Hashtag Activism) • Message Analysis: understanding what is resonating with users.
Leadership Movement Entrepreneurs	• The mass of personalized messages. (the clear messages for collective actions are hard to find.)	• User Analysis: understanding who is participating and what kind of data resonate with the users of the social movement.

presented with the corresponding solutions. You can see the challenges/solutions that were addressed in this dissertation in Table 5.1.

2. The overarching issue framing this dissertation was: *With the data collection constraints and fast paced content generation on online social media platforms, how can we improve the success of a hashtag activism by generating more engaging content for specific user type in real time?*

   (a) A real-time analytical framework was developed to identify user type of a hashtag campaign as individual (male/female) and organization, then the features salience showed that each user type was better identified by which feature set.

(b) A real-time analytical framework was developed to determine if a tweet is likely to be retweeted and spread in the given hashtag activism campaigns, then it analyzed which features from the content including fine grained and coarse grained features such as cluster of hashtags, topics and sentiment of the tweet were more likely to be retweeted in two engineering diversity related hashtag activism campaigns.

A summary of the findings with respect to the primary research questions is presented in Table 5.2.

5.1.1 Research Question 1: Real-Time User Type Classification

One of the main components of a social movements are actors participating in a shared cause as discussed before. Inspired by manual analysis of an engineering diversity related hashtag activism campaign, *ILookLikeAnEngineer*, identifying user type participation in online social campaign helped us understand the movement in a deeper level specially for triggers which helped the movement sustain and spread in the online world and tuning the messaging for the targeted audience. Hence, an analytical framework was devised to classify a Twitter user into three classes: organization, male and female in a real-time manner using different sets of features from a user and the content of the tweet. Although the primary incentive of such campaign was to create a role models for females and break the gender stereotypes in the engineering industry, the movements' messaging was found effective to engage more organizations in the course of action where their participation was more than 35% following women, 47%. It proves the hypothesis of first research question that diverse group of users (i.e. individuals with organizations altogether) helps a social

Table 5.2: List of research questions with their respective outcomes

Research Question	**Analysis**	**Outcomes** (*Solution Identified*)
What is the user type distribution participating in the campaign in real-time?	**Real-Time User Type Classification**	• Analytical framework to identify user type as individual(female/male) and organization in a unified real-time framework. (*Use of Hashtag, Machine Learning Algorithm, Real-Time Analysis*) • Image is a good discriminatory feature for organization whereas user name is a good predictor of gender for individuals(female/male). (*User Analysis*) • Given the fact that ILLAE was primarily targeted toward women, organizations had a substantial participation than expected (35% vs 10% on general). (*User Analysis*)
Is a tweet going to be engaging enough to help spread the campaign's message in real-time?	**Real-Time Retweetability Classification**	• Analytical framework to determine if a tweet is going to be retweeted in a given hashtag campaign based on the history of the campaign in real-time. (*Use of Hashtag, Machine Learning Algorithm, Real-Time Analysis*) • The feature set of content was shown more important than user's social network feature set. (*Message Analysis*) • Although the friends count is well correlated to the retweetability probability, monetary gain through advertisement were less likely to be propagated in the hashtag campaigns. (*Message Analysis*)

movement grow and spread in the online social network.

In terms of feature salience, it was shown that user profile pictures were a good predictor of the organization type whereas user name is good indicator of the individual type. It proves the fact that organizations tend to use non-human profile pictures in their accounts and users in these movements usually use their real names because they would like to have their voice heard and build an online collective identity.

This framework with the help of multi-modality characteristic of the features showed a good level of resilience toward the imbalance data which makes a better and more reliable tool for other hashtag activism campaigns compared to its counterparts that just rely on one type of data.

5.1.2 Research Question 2: Real-Time Retweetability Classification

A real-time analytical tool was devised to determine if the content of the message is attractive and relevant for the audience of a particular hashtag campaign (if it is going to be retweeted or not?). This framework helps to increase the level of exposure of the messaging in an online social campaign. It just relies on what is available in one tweet which makes it highly applicable to the online social platforms such as Twitter where there are restrictions imposed by the platform policies as one of the challenges of online campaigns and also allows activists and regular users with limited resource to spread their messages in the given movement efficiently .

This framework employed a set of fine grained features (*e.g. length of the tweet, user's network metadata*) and coarse grained features (*i.e. cluster of hashtags, topics and sentiments*). As the final results showed, sentiment (LIWC) was the most performant feature set. It proves the hypothesis that emotional conversation is leading to persuasion that eventually triggers information dissemination. In the two engineering diversity campaigns, informal tone of the messages was found the most retweetable quality in terms of sentiment. Also monetary gain and advertisement were less likely to be propagated in the campaign as opposed to the STEM related topics were deemed more retweetable and engaging in terms of topics and cluster of hashtags.

Knowing what kind of content resonates with the audience of the movement helps with promoting its goals and makes the information diffusion more efficient and faster compared to other baselines in real-time.

5.2 Limitations and Future Works

The first framework can be applied in another general dataset or campaign as all of those features are available regardless of the nature of the content. But it was shown that it worked better for campaign data as people creating content are more authentic and honest about their identity on Twitter rather than people on a general dataset (CrowdFlower).

The second framework was a specifically designed for hashtag activism campaigns where multiple hashtags are used and they are used as one of the feature sets. Also with the new tweets generated in the campaign, sentiments, topics and hashtags will change overtime and the model should be retrained to increase the accuracy of the model.

Also these two frameworks use real-time information meaning extra information such as user's historical tweets is not required, using another source of information might improve these frameworks.

This work can be improved in different ways. User type framework can be extended by including other user information like location, background profile image, the color of the user profiles and other elements of the tweets such as hashtags, URLs and media (photos and videos). For retweetability framework, incorporating latent features of media, photo and video along with the tweet as well as the analysis of hyperlinks is definitely an interesting research direction since one of the important features in activisms are media inclusion.

Also to extend the idea of increasing the resharing likelihood with attractive and relevant tweet, a recommender framework can be another efficient tool to recommend highly probable tweets to a user based on his social network metadata and his recent tweets in the given activism.

Although both frameworks used SVM as their classification algorithm due to it non-linearity capability, using deep neural network such as the work in [Zhang et al., 2016] as classification frameworks can lead to more accurate predictive capability.

5.3 Applications

As mentioned before, the first analytical framework can be applied to any other online activism including hashtag activism campaigns and other online social media platforms whereas the second framework was framed to utilize the characteristics of hashtag clusters, hence it could be applied to other hashtag activism campaigns.

In the first part of this work, real-time user type classification, although the results from this framework are not completely accurate, it gives a good sense of user type distribution in an online campaign in a much faster, more economical and efficient manner. Therefore, it is a great tool for real-time monitoring of user type distribution participating during the campaign course of action and helps people specially influencers of the campaign to tune their messages. It also could be applied when understanding the initial phases of these campaigns is important as long as the model was trained by another similar campaign. This way, we can have a faster and more accurate messaging strategy to target the desired user type at the early stage of campaign development.

The second framework provides a real-time tool for those participating in a hashtag-driven campaigns to increase the likelihood of the content exposure by predicting the retweetability of a tweet beforehand then they can tune the content of the tweet to make it more likely to be retweeted. As it relies on the content of the messages specially the cluster of hashtags, it should be applied to hashtag driven campaigns. Although it cannot be used for the initial phase of the campaign development because it needs some messaging history to build the model, a model trained in one social media platform, such as Twitter, could be applied to another one, such as Instagram, is an option for this framework to be used.

Moreover these two frameworks are applicable to other domains where identifying organizations, for instance emergency management, is crucial. In such situations, these frameworks, user type and retweetability classification, become more critical where time and resources are very crucial and limited since Twitter API limitation policies have made it challenging to respond promptly unless extra information is not needed. Also they can be utilized in online movements where identifying female/male participants (any campaign directed towards diversity of issues such as gender-based violence) and efficient messaging to the desired group of recipients are important and pivotal to the success of the campaign.

Appendix A: Model Selection

A.1 Performance Measures

In [Japkowicz and Shah, 2011], the following diagram, figure A.1, shows all kinds of performance measures according to the type of the data and classification algorithm. As you can see in the following figure, my problem is a multi-class classification problem that accuracy is one of the options for performance measure of the model.

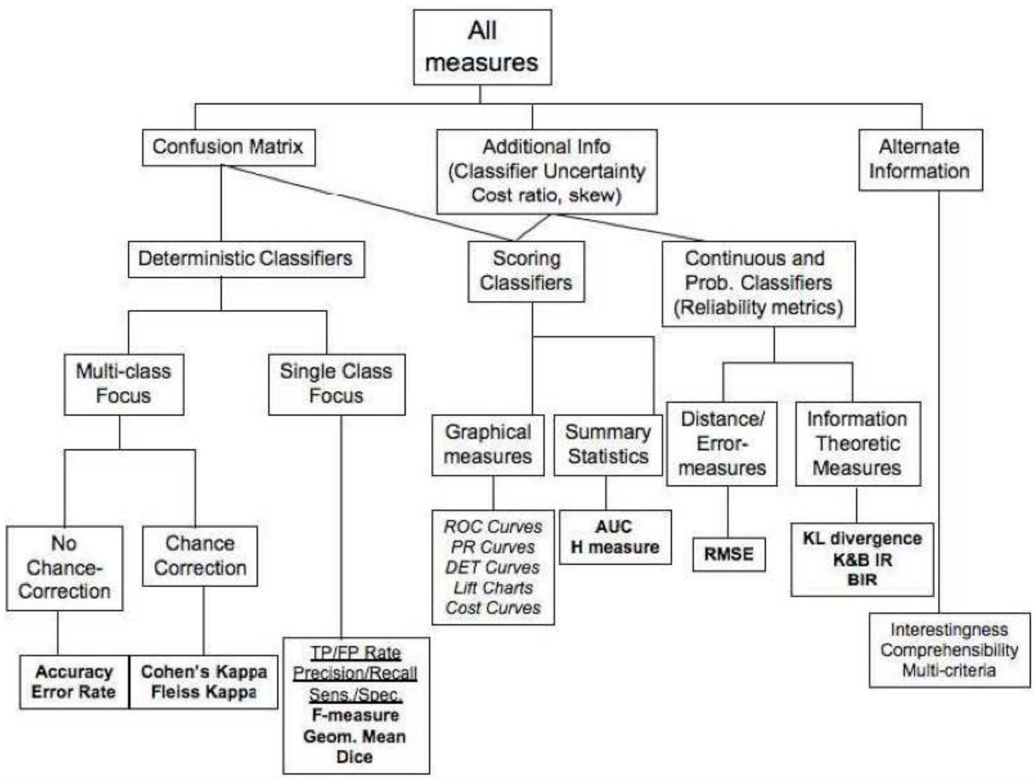

Figure A.1: Performance Measures

A.2 Model Evaluation

Stratification: We have to keep in mind that our dataset represents a random sample drawn from a probability distribution; and we typically assume that this sample is representative of the true population – more or less. Now, further subsampling without replacement alters the statistic (mean, proportion, and variance) of the sample. Hence, when splitting the data to different subsets of training or test, we have to make sure the statistics of each subset is the same as the entire dataset. This process is called **stratification** which makes sure that model evaluation is statistically sound.

A.2.1 Training-Test

The entire dataset is split into a separate training and test set with one hyperparameter setting (figure A.2). The model is trained on the training set the prespecified hyperparameters and then it will be evaluated on the test set based on the performance measure selected in the previous section. In this case, hyperparemeter optimization is not of concern and this method does not show the real performance of the model, so it is not very accurate.

Figure A.2: Training-Test split

A.2.2 Training-Validation-Test

In this model evaluation, hyperparameter optimization becomes important and we would like to find the optimal hyperparameters to evaluate the final model. So the entire dataset is split into a separate training, validation and test sets. Then by using training set, a model with a set of hyperparamters is trained and the evaluated by validation set. This process is reiterated few times to find the optimal hyperparameters that yields a maximum performance. Then a model with the optimal hyperparameters is trained over training and validation sets and it will be finally evaluated by the independent test set to get an unbiased estimate of its performance.

Figure A.3: Training-Validation-Test split

A.2.3 Holdout Cross Validation

The holdout k-fold cross validation is as follows:

1. A separate test set is smapled from the data and put aside for final testing.
2. For each of the k subsets of the remaining data set D, create a training set T=D-k.
3. Train the model on T with a set of hyperparameters.
4. Test the model on k.

5. Repeat the process for the remaining *k-1* partitions such that the next partition is treated as test and the rest will be used as training.

6. Overall accuracy of the model for that specific hyperparameter is averaged across all *k*-partitions.

7. Another set of hyperparameters will be selected by a selection methods; *i.e. random search, grid search or bayesian*.

8. The hyperparameters of the model with highest overall accuracy will be chosen for the final model. Train the model on D with the optimal hyperparameters.

9. Test the model with the test set.

10. The performance measure of the model is calculated for model selection explained in the next section.

So the following figures A.4 and A.5 show the process of this method

Figure A.4: The overall picture of the k-fold CV

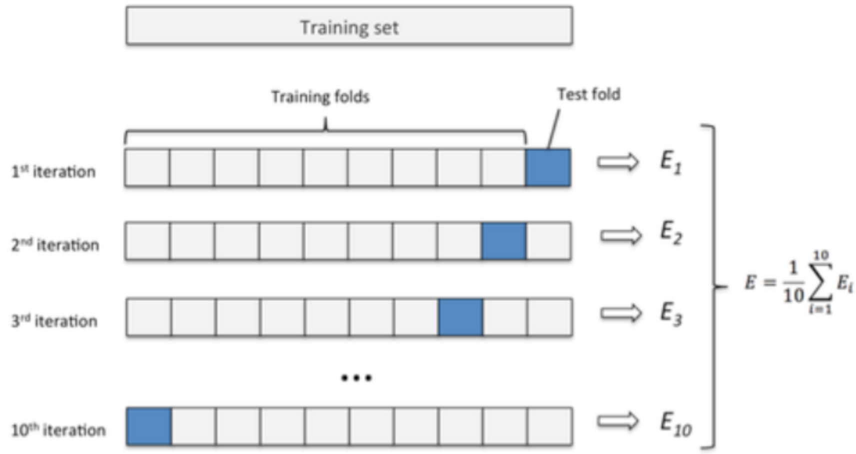

Figure A.5: The model with selected hyperparameters

A.2.4 Repeated Holdout Cross Validation

To have a more accurate model evaluation over the dataset, it is suggested to rerun the holdout CV multiple times and the final evaluation measure of the model will be the statistics of the repeated runs. This method gives a chance to more data to be in a test set and gives more accuracy but it lacks the idea of giving each data instance a chance to be tested.

A.2.5 Nested Cross Validation

To overcome the shortcoming mentioned in the repeated holdout CV, nested CV suggests to have two folds; *outer fold* and *inner fold*. Using nested cross-validation the model will be trained for different k fold, 1 for each of the k outer folds, and the inner folds are used to optimize the hyperparameters of each model (e.g., using gridsearch in combination with another k-fold cross-validation. If your model is stable, these k models should all have the same hyperparameter values, and the average performance of this model is the evalutaion of the model based on the outer test folds. If the results of the outer CV for optimal hyperparameters differs a lot between folds, then this means that you cannot be sure which hyperparameters set will generalize well over the entire dataset. To perform the final hyperparameter tuning, you do the inner loop on the whole data and pick the best according to the same criterion you used in the nested CV. Then you fit it to the whole data and that's your final model, but the predictive performance of your model is evaluated by the nested CV.

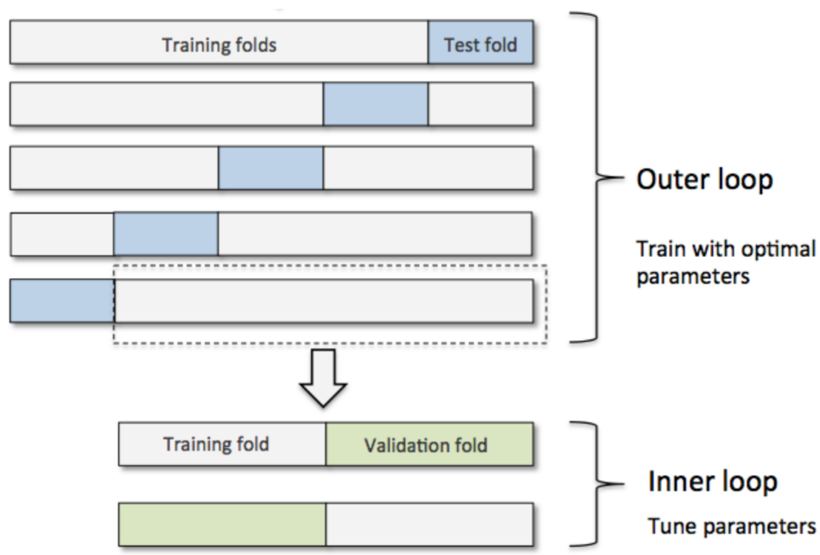

Figure A.6: The Nested CV

Appendix B: Statistical Comparison

Comparing machine learning methods and selecting a final model is a common operation in applied machine learning. Models are commonly evaluated using one of the methods explained in the previous secion like k-fold cross-validation from which mean the performance measures are calculated and compared directly. Although simple, this approach can be misleading as it is hard to know whether the difference between the averaged measures is real or the result of a statistical chance.

Statistical significance tests are designed to address this problem and quantify the likelihood of the samples of measures being observed given the assumption that they were drawn from the same distribution. If this assumption, or null hypothesis, is rejected, it suggests that two models are statistically significant and we can choose a more optimal model at the end.

B.1 Statistical Hypothesis Tests

Generally, a statistical hypothesis test for comparing samples quantifies how likely it is to observe two data samples given the assumption that the samples have the same distribution. The assumption of a statistical test is called the null hypothesis and we can calculate statistical measures and interpret them in order to decide whether or not to accept or reject the null hypothesis.

In the case of selecting models based on their performance measures, we are interested to know whether there is a real or statistically significant difference between the two models. If the result of the test suggests that there is insufficient evidence to reject the null hypothesis, then any observed difference in model skill is likely due to statistical chance.

If the result of the test suggests that there is sufficient evidence to reject the null hypothesis, then any observed difference in model skill is likely due to a difference in the models. So there are various ways to run the test.

B.2 Comparisons of Classifiers

The following tests are the summary of various methods and tests for model selection from [Dietterich, 1998] and [Demšar, 2006].

B.2.1 Averaging

Machine learning papers compute the average classification accuracies of classifiers across the tested data sets. Averages are susceptible to outliers. They allow classifier's excellent performance on one data set to compensate for the overall bad performance, or the opposite, a total failure on one domain can prevail over the fair results on most others. There may be situations in which such behaviour is desired, while in general we probably prefer classifiers that behave well on as many problems as possible, which makes averaging over data sets inappropriate. Also averages are also not used (nor useful) for statistical tests.

B.2.2 Paired T-Test

It is common practice to evaluate classification methods using classification accuracy, to evaluate each model using k-fold cross-validation, to assume a Gaussian distribution for the sample of k models' performance measures, and to use the mean of the sample as a summary of the model's measure. We could require that each classifier evaluated using this procedure be evaluated on exactly the same splits of the dataset via k-fold cross-validation. This would give samples of matched paired measures between two classifiers, matched because each classifier was evaluated on the same k test sets.

Then the paired Student's t-test will be used to check if the difference in the mean accuracy between the two models is statistically significant, e.g. reject the null hypothesis that assumes that the two samples have the same distribution. In fact, this is a common way to compare classifiers with perhaps hundreds of published papers using this methodology. The problem is, a key assumption of the paired Student's t-test has been violated. Namely, the observations in each sample are not independent. As part of the k-fold cross-validation procedure, a given observation will be used in the training dataset (k-1) times.

This means that the estimated performance measure are dependent, not independent, and in turn that the calculation of the t-statistic in the test will be misleadingly wrong along with any interpretations of the statistic and p-value. This observation requires a careful understanding of both the resampling method used, in this case k-fold cross-validation, and the expectations of the chosen hypothesis test, in this case the paired Student's t-test. Without this background, the test appears appropriate, a result will be calculated and interpreted, and everything will look fine.

B.2.3 Wilcoxon Signed-Ranks Test

The Wilcoxon signed-ranks test ,[Wilcoxon, 1945], is a non-parametric alternative to the paired t-test, which ranks the differences in performances of two classifiers for each data set, ignoring the signs, and compares the ranks for the positive and the negative differences. This approach makes fewer assumptions, such as not assuming that the distribution of the performance measures is normal distribution. Although the test is nonparametric, it still assumes that the observations within each sample are independent (e.g. iid), and using k-fold cross-validation would create dependent samples and violate this assumption.

B.2.4 McNemar Test

The corrected version of paired T-test is McNemar's test. In [Dietterich, 1998] it is recommended to use the McNemar's statistical hypothesis test in cases where there is a limited amount of data and each algorithm can only be evaluated once. McNemar's test is like the Chi-Squared test, and in this case is used to determine whether the difference in observed proportions in the algorithm's confusion matrix, figure B.1, are significantly different from the expected proportions.

To apply McNemar's test, [Everitt, 1977], the confusion matrix like figure B.1 from the test results of the two models is constructed then by using the equation (B.1), under the null hypothesis, with a sufficiently large number of difference; i.e. $(e_{10} + e_{01}) \geq 20$, χ^2 has a chi-squared distribution with 1 degree of freedom. If the χ^2 result is significant, this provides sufficient evidence to reject the null hypothesis, in favour of the alternative hypothesis that the two models are different.

Algo 1 / Algo 2	right	wrong
right	number of examples well classified by both	e_{01} number of examples well classified by 1 but not by 2
wrong	e_{10} number of examples missclassified by 1 but not by 2	number of examples missclassified by both

Figure B.1: McNemar Confusion Matrix

$$\chi^2 = \frac{(\mid e_{10} - e_{01} \mid -1)^2}{(e_{10} + e_{01})} \tag{B.1}$$

B.2.5 Friedman Test

This test is used when multiple classifiers over multiple datasets are to be compared. This test ensures that overall which classifiers are different from each other. The requirement for this test is at least the result of test for five models over 10 different datasets is available. It uses a ranking method of models and F-distribution, [Demšar, 2006].

Bibliography

Bibliography

[Nat, 2007] (2007). National science foundation (2007) women, minorities, and persons with disabilities in science and engineering. arlington, va: Division of science resources statistics.

[ASE, 2009] (2009). American society of engineering education.

[UKR, 2010] (2010). Uk resource centre for women in science, engineering and technology (2010) students in higher education 2009.

[Alowibdi et al., 2013] Alowibdi, J. S., Buy, U. A., and Yu, P. (2013). Language independent gender classification on twitter. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 739–743. ACM.

[Argamon et al., 2007] Argamon, S., Koppel, M., Pennebaker, J. W., and Schler, J. (2007). Mining the blogosphere: Age, gender and the varieties of self-expression. *First Monday*, 12(9).

[Bao et al., 2013] Bao, P., Shen, H.-W., Chen, W., and Cheng, X.-Q. (2013). Cumulative effect in information diffusion: empirical study on a microblogging network. *PloS one*, 8(10):e76027.

[Bastalich et al., 2003] Bastalich, W., Mills, J., Franzway, S., Gill, J., Sharp, R., et al. (2003). 'i had this real feeling that it was a boys club'. In *Engineering Education for a Sustainable Future: Proceedings of the 14th Annual Conference for Australasian Association for Engineering Education and 9th Australasian Women in Engineering Forum*, page 499. Australasian Association for Engineering Education.

[Bastos et al., 2014] Bastos, M. T., da Cunha Recuero, R., and da Silva Zago, G. (2014). Taking tweets to the streets: A spatial analysis of the vinegar protests in brazil. *First Monday*, 19(3).

[Bay et al., 2006] Bay, H., Tuytelaars, T., and Van Gool, L. (2006). Surf: Speeded up robust features. *Computer vision–ECCV 2006*, pages 404–417.

[Bayer et al., 2012] Bayer, M., Sommer, W., and Schacht, A. (2012). Font size matters—emotion and attention in cortical responses to written words. *PloS one*, 7(5):e36042.

[Bennett and Segerberg, 2012] Bennett, W. L. and Segerberg, A. (2012). The logic of connective action: Digital media and the personalization of contentious politics. *Information, communication & society*, 15(5):739–768.

[Bergsma et al., 2013] Bergsma, S., Dredze, M., Van Durme, B., Wilson, T., and Yarowsky, D. (2013). Broadly improving user classification via communication-based name and location clustering on twitter. In *HLT-NAACL*, pages 1010–1019.

[Blei et al., 2003] Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.

[Borge-Holthoefer et al., 2015] Borge-Holthoefer, J., Magdy, W., Darwish, K., and Weber, I. (2015). Content and network dynamics behind egyptian political polarization on twitter. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 700–711. ACM.

[Boyce et al., 2002] Boyce, M. C., Chisholm, P., Crawley, E., Gibson, L., Gleason, K., Lynch, N., and Vander Sande, J. (2002). Report of the committee on women faculty in the school of engineering at mit.

[Braccini et al., 2019] Braccini, A. M., Sæbø, Ø., and Federici, T. (2019). From the blogosphere into the parliament: The role of digital technologies in organizing social movements. *Information and Organization*.

[Burger et al., 2011] Burger, J. D., Henderson, J., Kim, G., and Zarrella, G. (2011). Discriminating gender on twitter. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1301–1309. Association for Computational Linguistics.

[Burger and Henderson, 2006] Burger, J. D. and Henderson, J. C. (2006). An exploration of observable features related to blogger age. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs*, pages 15–20. Menlo Park, CA.

[Cardoso et al., 2019] Cardoso, A., Boudreau, M.-C., and Carvalho, J. Á. (2019). Organizing collective action: Does information and communication technology matter? *Information and Organization*, page 100256.

[Chen et al., 2009] Chen, W., Wang, Y., and Yang, S. (2009). Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 199–208.

[Chen et al., 2015] Chen, X., Wang, Y., Agichtein, E., and Wang, F. (2015). A comparative study of demographic attribute inference in twitter. *ICWSM*, 15:590–593.

[Cheng et al., 2014] Cheng, J., Adamic, L., Dow, P. A., Kleinberg, J. M., and Leskovec, J. (2014). Can cascades be predicted? In *Proceedings of the 23rd international conference on World wide web*, pages 925–936. ACM.

[Cheng et al., 2018] Cheng, J., Kleinberg, J., Leskovec, J., Liben-Nowell, D., Subbian, K., Adamic, L., et al. (2018). Do diffusion protocols govern cascade growth? In *Twelfth International AAAI Conference on Web and Social Media*.

[Christensen, 2011] Christensen, H. S. (2011). Political activities on the internet: Slacktivism or political participation by other means? *First Monday*, 16(2).

[Chung and Pennebaker, 2007] Chung, C. and Pennebaker, J. W. (2007). The psychological functions of function words. *Social communication*, pages 343–359.

[Chung, 2017] Chung, J. E. (2017). Retweeting in health promotion: Analysis of tweets about breast cancer awareness month. *Computers in Human Behavior*, 74:112–119.

[Clark, 2016] Clark, R. (2016). "hope in a hashtag": the discursive activism of# whyistayed. *Feminist Media Studies*, 16(5):788–804.

[De Choudhury et al., 2012] De Choudhury, M., Diakopoulos, N., and Naaman, M. (2012). Unfolding the event landscape on twitter: classification and exploration of user categories. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, pages 241–244. ACM.

[De Choudhury et al., 2016] De Choudhury, M., Jhaver, S., Sugar, B., and Weber, I. (2016). Social media participation in an activist movement for racial equality. In *ICWSM*, pages 92–101.

[De Silva and Riloff, 2014] De Silva, L. and Riloff, E. (2014). User type classification of tweets with implications for event recognition. *ACL 2014*, page 98.

[Demšar, 2006] Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine learning research*, 7(Jan):1–30.

[Diani, 1992] Diani, M. (1992). The concept of social movement. *The sociological review*, 40(1):1–25.

[Dietterich, 1998] Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural computation*, 10(7):1895–1923.

[Domingos and Richardson, 2001] Domingos, P. and Richardson, M. (2001). Mining the network value of customers. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 57–66.

[Earl and Schussman, 2002] Earl, J. and Schussman, A. (2002). The new site of activism: On-line organizations, movement entrepreneurs, and the changing location of social movement decision making. In *Consensus decision making, Northern Ireland and indigenous movements*, pages 155–187. Emerald Group Publishing Limited.

[Eltantawy and Wiest, 2011] Eltantawy, N. and Wiest, J. B. (2011). The arab spring| social media in the egyptian revolution: reconsidering resource mobilization theory. *International journal of communication*, 5:18.

[Everitt, 1977] Everitt, B. (1977). The analysis of contingency tables.

[Evetts, 1998] Evetts, J. (1998). Managing the technology but not the organization: Women and career in engineering. *Women in management review*, 13(8):283–290.

[Gargi et al., 2011] Gargi, U., Lu, W., Mirrokni, V., and Yoon, S. (2011). Large-scale community detection on youtube for topic discovery and exploration. In *Fifth International AAAI Conference on Weblogs and Social Media*.

[Geman and Geman, 1987] Geman, S. and Geman, D. (1987). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. In *Readings in computer vision*, pages 564–584. Elsevier.

[George and Leidner, 2019] George, J. J. and Leidner, D. E. (2019). From clicktivism to hacktivism: Understanding digital activism. *Information and Organization*.

[Golbeck et al., 2010] Golbeck, J., Grimes, J. M., and Rogers, A. (2010). Twitter use by the us congress. *Journal of the Association for Information Science and Technology*, 61(8):1612–1621.

[Gonzalez and Kuenzi, 2012] Gonzalez, H. B. and Kuenzi, J. J. (2012). Science, technology, engineering, and mathematics (stem) education: A primer. Congressional Research Service, Library of Congress Washington, DC.

[Guerini et al., 2011] Guerini, M., Strapparava, C., and Ozbal, G. (2011). Exploring text virality in social networks. In *Fifth international AAAI conference on weblogs and social media*.

[Gunn, 2015] Gunn, C. (2015). Hashtagging from the margins: Women of color engaged in feminist consciousness-raising on twitter. *Women of Color and Social Media Multitasking: Blogs, Timelines, Feeds, and Community*, pages 21–34.

[Gurr, 2015] Gurr, T. R. (2015). *Why men rebel*. Routledge.

[Hänska Ahy, 2016] Hänska Ahy, M. (2016). Networked communication and the arab spring: Linking broadcast and social media. *New Media & Society*, 18(1):99–116.

[He et al., 2015] He, J., Hong, L., Frias-Martinez, V., and Torrens, P. (2015). Uncovering social media reaction pattern to protest events: a spatiotemporal dynamics perspective of ferguson unrest. In *International Conference on Social Informatics*, pages 67–81. Springer.

[Heath, 1996] Heath, C. (1996). Do people prefer to pass along good or bad news? valence and relevance of news as predictors of transmission propensity. *Organizational behavior and human decision processes*, 68(2):79–94.

[Higgs, 2015] Higgs, E. T. (2015). # justiceforliz: Power and privilege in digital transnational women's rights activism. *Feminist Media Studies*, 15(2):344–347.

[Highfield, 2017] Highfield, T. (2017). *Social media and everyday politics*. John Wiley & Sons.

[Hopke, 2015] Hopke, J. E. (2015). Hashtagging politics: Transnational anti-fracking movement twitter practices. *Social Media+ Society*, 1(2):2056305115605521.

[Hosmer Jr et al., 2013] Hosmer Jr, D. W., Lemeshow, S., and Sturdivant, R. X. (2013). *Applied logistic regression*, volume 398. John Wiley & Sons.

[Huang and Xue, 2015] Huang, J. and Xue, J. (2015). The predictive power of content and temporal features of posts in information dissemination in microblogging. *Journal of China Tourism Research*, 11(2):150–165.

[Huggins, 2002] Huggins, R. (2002). Democracy and. *Politics: An Introduction*, 16:159.

[James and Van Seters, 2014] James, P. and Van Seters, P. (2014). Global social movements and global civil society: A critical overview. *Globalization and Politics, Vol. 2: Global Social Movements and Global Civil Society*, pages vii–xxx.

[Japkowicz and Shah, 2011] Japkowicz, N. and Shah, M. (2011). *Evaluating learning algorithms: a classification perspective*. Cambridge University Press.

[Johri et al., 2018a] Johri, A., Heyman-Schrum, C., Ruiz, D., Malik, A., Karbasian, H., Handa, R., and Purohit, H. (2018a). More than an engineer: Intersectional self-expressions in a hashtag activism campaign for engineering diversity. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, page 44. ACM.

[Johri et al., 2018b] Johri, A., Karbasian, H., Malik, A., Handa, R., and Purohit, H. (2018b). How diverse users and activities trigger connective action via social media: Lessons from the twitter hashtag campaign# ilooklikeanengineer. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.

[Jones, 1997] Jones, K. S. (1997). *Readings in information retrieval*. Morgan Kaufmann.

[Joyce, 2010] Joyce, M. C. (2010). *Digital activism decoded: The new mechanics of change*. IDEA.

[Kameda, 2011] Kameda, A. (2011). The advancement of women in science and technology. *Transforming Japan: How Feminism and Diversity are Making a Difference, The Feminist Press, New York, NY*, pages 89–101.

[Karbasian et al., 2018] Karbasian, H., Purohit, H., Handa, R., Malik, A., and Johri, A. (2018). Real-time inference of user types to assist with more inclusive and diverse social media activism campaigns. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, pages 171–177. ACM.

[Karpf, 2010] Karpf, D. (2010). Online political mobilization from the advocacy group's perspective: Looking beyond clicktivism. *Policy & Internet*, 2(4):7–41.

[Kempe et al., 2003] Kempe, D., Kleinberg, J., and Tardos, É. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146.

[Khondker, 2011] Khondker, H. H. (2011). Role of the new media in the arab spring. *Globalizations*, 8(5):675–679.

[Kissler et al., 2007] Kissler, J., Herbert, C., Peyk, P., and Junghofer, M. (2007). Buzzwords: early cortical responses to emotional words during reading. *Psychological Science*, 18(6):475–480.

[Kotsenas et al., 2018] Kotsenas, A. L., Arce, M., Aase, L., Timimi, F. K., Young, C., and Wald, J. T. (2018). The strategic imperative for the use of social media in health care. *Journal of the American College of Radiology*, 15(1):155–161.

[Kupavskii et al., 2012] Kupavskii, A., Ostroumova, L., Umnov, A., Usachev, S., Serdyukov, P., Gusev, G., and Kustarev, A. (2012). Prediction of retweet cascade size over time. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, pages 2335–2338. ACM.

[Lee and Hsieh, 2013] Lee, Y.-H. and Hsieh, G. (2013). Does slacktivism hurt activism?: the effects of moral balancing and consistency in online activism. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 811–820. ACM.

[Lewis et al., 2007] Lewis, S., Harris, R., Cox, B., et al. (2007). Engineering a better workplace: A diversity guide for the engineering profession. *Engineering a Better Workplace: a Diversity Guide for the Engineering Profession*, page xviii.

[Li et al., 2016] Li, Y., Liu, T., Jiang, J., and Zhang, L. (2016). Hashtag recommendation with topical attention-based lstm. Coling.

[Liu and Ruths, 2013] Liu, W. and Ruths, D. (2013). What's in a name? using first names as features for gender inference in twitter. In *AAAI spring symposium: Analyzing microtext*, volume 13, page 01.

[Lowe, 2004] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110.

[Luminet IV et al., 2000] Luminet IV, O., Bouts, P., Delie, F., Manstead, A. S., and Rimé, B. (2000). Social sharing of emotion following exposure to a negatively valenced situation. *Cognition & Emotion*, 14(5):661–688.

[Luo et al., 2013] Luo, Z., Osborne, M., Tang, J., and Wang, T. (2013). Who will retweet me?: finding retweeters in twitter. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, pages 869–872. ACM.

[Malcom-Piqueux and Malcom, 2013] Malcom-Piqueux, L. and Malcom, S. M. (2013). Engineering diversity: Fixing the educational system to promote equity. *Bridge*, 43(1):24–34.

[Malik et al., 2018] Malik, A., Johri, A., Handa, R., Karbasian, H., and Purohit, H. (2018). # ilooklikeanengineer: Using social media based hashtag activism campaigns as a lens to better understand engineering diversity issues. *arXiv preprint arXiv:1805.01971*.

[McCallum, 2002] McCallum, A. K. (2002). Mallet: A machine learning for language toolkit.

[McCarthy and Zald, 1977] McCarthy, J. D. and Zald, M. N. (1977). Resource mobilization and social movements: A partial theory. *American journal of sociology*, 82(6):1212–1241.

[McCaughey and Ayers, 2013] McCaughey, M. and Ayers, M. D. (2013). *Cyberactivism: Online activism in theory and practice*. Routledge.

[McCorriston et al., 2015] McCorriston, J., Jurgens, D., and Ruths, D. (2015). Organizations are users too: Characterizing and detecting the presence of organizations on twitter. In *Proceedings of the 9th International AAAI Conference on Weblogs and Social Media (ICWSM)*, pages 650–653.

[McIlwee and Robinson, 1992] McIlwee, J. S. and Robinson, J. G. (1992). *Women in engineering: Gender, power, and workplace culture*. SUNY Press.

[McNemar, 1947] McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157.

[Milbrath and Goel, 1977] Milbrath, L. W. and Goel, M. L. (1977). *Political participation: How and why do people get involved in politics?* Rand McNally College Pub. Co.

[Miller, 2004] Miller, G. E. (2004). Frontier masculinity in the oil industry: The experience of women engineers. *Gender, Work & Organization*, 11(1):47–73.

[Minkus et al., 2015] Minkus, T., Liu, K., and Ross, K. W. (2015). Children seen but not heard: When parents compromise children's online privacy. In *Proceedings of the 24th International Conference on World Wide Web*, pages 776–786. International World Wide Web Conferences Steering Committee.

[Mukherjee and Liu, 2010] Mukherjee, A. and Liu, B. (2010). Improving gender classification of blog authors. In *Proceedings of the 2010 conference on Empirical Methods in natural Language Processing*, pages 207–217. Association for Computational Linguistics.

[Mulac et al., 1986] Mulac, A., Lundell, T. L., and Bradac, J. J. (1986). Male/female language differences and attributional consequences in a public speaking situation: Toward an explanation of the gender-linked language effect. *Communications Monographs*, 53(2):115–129.

[Mulac et al., 1990] Mulac, A., Studley, L. B., and Blau, S. (1990). The gender-linked language effect in primary and secondary students' impromptu essays. *Sex Roles*, 23(9-10):439–470.

[Nan and Lu, 2014] Nan, N. and Lu, Y. (2014). Harnessing the power of self-organization in an online community during organizational crisis. *Mis Quarterly*, 38(4):1135–1157.

[Naveed et al., 2011] Naveed, N., Gottron, T., Kunegis, J., and Alhadi, A. C. (2011). Bad news travel fast: A content-based analysis of interestingness on twitter. In *Proceedings of the 3rd international web science conference*, page 8. ACM.

[Neppalli et al., 2016] Neppalli, V. K., Medeiros, M. C., Caragea, C., Caragea, D., Tapia, A. H., and Halse, S. E. (2016). Retweetability analysis and prediction during hurricane sandy. In *ISCRAM*.

[Newman et al., 2008] Newman, M. L., Groom, C. J., Handelman, L. D., and Pennebaker, J. W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes*, 45(3):211–236.

[Oberschall, 1973] Oberschall, A. (1973). *Social conflict and social movements*. Prentice hall.

[Palermo, 2004] Palermo, J. (2004). Breaking the cultural mould: the key to womens career success.

[Papacharissi and de Fatima Oliveira, 2012] Papacharissi, Z. and de Fatima Oliveira, M. (2012). Affective news and networked publics: The rhythms of news storytelling on# egypt. *Journal of communication*, 62(2):266–282.

[Pennebaker et al., 2015] Pennebaker, J. W., Boyd, R. L., Jordan, K., and Blackburn, K. (2015). The development and psychometric properties of liwc2015. Technical report.

[Pennebaker et al., 2003] Pennebaker, J. W., Mehl, M. R., and Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, 54(1):547–577.

[Pervin et al., 2014] Pervin, N., Takeda, H., and Toriumi, F. (2014). Factors affecting retweetability: an event-centric analysis on twitter. *ICIS 2014 Proceedings*.

[Peters et al., 2009] Peters, K., Kashima, Y., and Clark, A. (2009). Talking about others: Emotionality and the dissemination of social information. *European Journal of Social Psychology*, 39(2):207–222.

[Petrovic et al., 2011] Petrovic, S., Osborne, M., and Lavrenko, V. (2011). Rt to win! predicting message propagation in twitter. In *Fifth International AAAI Conference on Weblogs and Social Media*.

[Phipps, 2002] Phipps, A. (2002). Engineering women: Thegendering'of professional identities. *International Journal of Engineering Education*, 18(4):409–414.

[Powell et al., 2009] Powell, A., Bagilhole, B., and Dainty, A. (2009). How women engineers do and undo gender: Consequences for gender equality. *Gender, work & organization*, 16(4):411–428.

[Preece, 2000] Preece, J. (2000). Online communities: designing usability, supporting sociability. *Industrial Management & Data Systems*, 100(9):459–460.

[Preece, 2001] Preece, J. (2001). Sociability and usability in online communities: Determining and measuring success. *Behaviour & Information Technology*, 20(5):347–356.

[Purohit et al., 2011] Purohit, H., Ruan, Y., Joshi, A., Parthasarathy, S., and Sheth, A. (2011). Understanding user-community engagement by multi-faceted features: A case study on twitter. In *WWW 2011 Workshop on Social Media Engagement (SoME)*.

[Radley and Bell, 2007] Radley, A. and Bell, S. E. (2007). Artworks, collective experience and claims for social justice: The case of women living with breast cancer. *Sociology of health & illness*, 29(3):366–390.

[Rainie et al., 2012] Rainie, L., Smith, A., Schlozman, K. L., Brady, H., and Verba, S. (2012). Social media and political engagement. *Pew Internet & American Life Project*, 19:2–13.

[Reuter and Kaufhold, 2018] Reuter, C. and Kaufhold, M.-A. (2018). Fifteen years of social media in emergencies: a retrospective review and future directions for crisis informatics. *Journal of Contingencies and Crisis Management*, 26(1):41–57.

[Richardson and Domingos, 2002] Richardson, M. and Domingos, P. (2002). Mining knowledge-sharing sites for viral marketing. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 61–70.

[Rickford, 2016] Rickford, R. (2016). Black lives matter: Toward a modern practice of mass struggle. In *New Labor Forum*, volume 25, pages 34–42. SAGE Publications Sage CA: Los Angeles, CA.

[Rimé, 2009] Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion review*, 1(1):60–85.

[Roberts and Ayre, 2002] Roberts, P. and Ayre, M. (2002). Did she jump or was she pushed? a study of women's retention in the engineering workforce. *International Journal of Engineering Education*, 18(4):415–421.

[Rucht, 1991] Rucht, D. (1991). *Research on social movements: the state of the art in Western Europe and the USA*. Westview Pr.

[Rudat and Buder, 2015] Rudat, A. and Buder, J. (2015). Making retweeting social: The influence of content and context information on sharing news in twitter. *Computers in Human Behavior*, 46:75–84.

[Rudat et al., 2014] Rudat, A., Buder, J., and Hesse, F. W. (2014). Audience design in twitter: Retweeting behavior between informational value and followers' interests. *Computers in human behavior*, 35:132–139.

[Rupp and Taylor, 1999] Rupp, L. J. and Taylor, V. (1999). Forging feminist identity in an international movement: a collective identity approach to twentieth-century feminism. *Signs: Journal of Women in Culture and Society*, 24(2):363–386.

[Saleem et al., 2014] Saleem, H. M., Xu, Y., and Ruths, D. (2014). Novel situational information in mass emergencies: What does twitter provide? *Procedia Engineering*, 78:155–164.

[Sandoval-Almazan and Gil-Garcia, 2014] Sandoval-Almazan, R. and Gil-Garcia, J. R. (2014). Towards cyberactivism 2.0? understanding the use of social media and other information technologies for political activism and social movements. *Government Information Quarterly*, 31(3):365–378.

[Schäfer, 2006] Schäfer, A. I. (2006). A new approach to increasing diversity in engineering at the example of women in engineering. *European Journal of Engineering Education*, 31(6):661–671.

[Schler et al., 2006] Schler, J., Koppel, M., Argamon, S., and Pennebaker, J. W. (2006). Effects of age and gender on blogging. In *AAAI spring symposium: Computational approaches to analyzing weblogs*, volume 6, pages 199–205.

[Schradie, 2018] Schradie, J. (2018). The digital activism gap: How class and costs shape online collective action. *Social Problems*, 65(1):51–74.

[Schussman and Earl, 2004] Schussman, A. and Earl, J. (2004). From barricades to firewalls? strategic voting and social movement leadership in the internet age. *Sociological Inquiry*, 74(4):439–463.

[Schwartz et al., 2013a] Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Lucas, R. E., Agrawal, M., Park, G. J., Lakshmikanth, S. K., Jha, S., Seligman, M. E., et al. (2013a). Characterizing geographic variation in well-being using tweets. In *ICWSM*, pages 583–591.

[Schwartz et al., 2013b] Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M. E., et al. (2013b). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, 8(9):e73791.

[Segerberg and Bennett, 2011] Segerberg, A. and Bennett, W. L. (2011). Social media and the organization of collective action: Using twitter to explore the ecologies of two climate change protests. *The Communication Review*, 14(3):197–215.

[Selander and Jarvenpaa, 2016] Selander, L. and Jarvenpaa, S. L. (2016). Digital action repertoires and transforming a social movement organization. *mis Quarterly*, 40(2):331–352.

[Simonyan and Zisserman, 2014] Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[Skoric, 2012] Skoric, M. M. (2012). What is slack about slacktivism. *Methodological and conceptual issues in cyber activism research*, 77:77–92.

[Smelser, 1963] Smelser, N. (1963). Theory of collective behavior. *New York: Free Press*.

[Smith and Petty, 1996] Smith, S. M. and Petty, R. E. (1996). Message framing and persuasion: A message processing analysis. *Personality and Social Psychology Bulletin*, 22(3):257–268.

[Son et al., 2013] Son, I., Lee, D., and Kim, Y. (2013). Understanding the effect of message content and user identity on information diffusion in online social networks. In *PACIS*, page 8.

[Stewart and Schultze, 2019] Stewart, M. and Schultze, U. (2019). Producing solidarity in social media activism: The case of my stealthy freedom. *Information and Organization*.

[Stieglitz and Dang-Xuan, 2013] Stieglitz, S. and Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of management information systems*, 29(4):217–248.

[Suh et al., 2010] Suh, B., Hong, L., Pirolli, P., and Chi, E. H. (2010). Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. In *2010 IEEE Second International Conference on Social Computing*, pages 177–184. IEEE.

[Sutton et al., 2015a] Sutton, J., Gibson, C. B., Phillips, N. E., Spiro, E. S., League, C., Johnson, B., Fitzhugh, S. M., and Butts, C. T. (2015a). A cross-hazard analysis of terse message retransmission on twitter. *Proceedings of the National Academy of Sciences*, 112(48):14793–14798.

[Sutton et al., 2015b] Sutton, J., Gibson, C. B., Spiro, E. S., League, C., Fitzhugh, S. M., and Butts, C. T. (2015b). What it takes to get passed on: message content, style, and structure as predictors of retransmission in the boston marathon bombing response. *PLoS one*, 10(8):e0134452.

[Tan et al., 2002] Tan, C.-M., Wang, Y.-F., and Lee, C.-D. (2002). The use of bigrams to enhance text categorization. *Information processing & management*, 38(4):529–546.

[Tan et al., 2013] Tan, L., Ponnam, S., Gillham, P., Edwards, B., and Johnson, E. (2013). Analyzing the impact of social media on social movements: a computational study on twitter and the occupy wall street movement. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 1259–1266. ACM.

[Taylor et al., 2004] Taylor, V. et al. (2004). "get up, stand up": tactical repertoires of social movements. *The Blackwell companion to social movements*, pages 262–293.

[Taylor et al., 1995] Taylor, V., Whittier, N., et al. (1995). Analytical approaches to social movement culture: The culture of the women's movement. *Social movements and culture*, 4:163–187.

[Thomson and Murachver, 2001] Thomson, R. and Murachver, T. (2001). Predicting gender from electronic discourse. *British Journal of Social Psychology*, 40(2):193–208.

[Tietjen, 2004] Tietjen, J. S. (2004). Why so few women, still? *IEEE Spectrum*, 41(10):57–58.

[Tilly, 2006] Tilly, C. (2006). Wunc. *Crowds*, pages 289–306.

[Tilly, 2019] Tilly, C. (2019). *Social Movements, 1768-2004*. Routledge.

[Tumasjan et al., 2010] Tumasjan, A., Sprenger, T. O., Sandner, P. G., and Welpe, I. M. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. *Icwsm*, 10(1):178–185.

[Turner et al., 1957] Turner, R. H., Killian, L. M., et al. (1957). *Collective behavior*, volume 3. Prentice-Hall Englewood Cliffs, NJ.

[Vaast et al., 2017] Vaast, E., Safadi, H., Lapointe, L., and Negoita, B. (2017). Social media affordances for connective action-an examination of microblogging use during the gulf of mexico oil spill. *MIS Quarterly*.

[Van Dijck, 2013] Van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press.

[Van Laer and Van Aelst, 2010] Van Laer, J. and Van Aelst, P. (2010). Internet and social movement action repertoires: Opportunities and limitations. *Information, Communication & Society*, 13(8):1146–1171.

[Vapnik, 1995] Vapnik, V. (1995). The nature of statistical learning theory· 6·[mj new york. *Springer-Verlag*, 1:995.

[Vapnik and Vapnik, 1998] Vapnik, V. and Vapnik, V. (1998). Statistical learning theory.

[Vegh, 2013] Vegh, S. (2013). Classifying forms of online activism: The case of cyberprotests against the world bank. In *Cyberactivism*, pages 81–106. Routledge.

[Volkova et al., 2015] Volkova, S., Bachrach, Y., Armstrong, M., and Sharma, V. (2015). Inferring latent user properties from texts published in social media. In *AAAI*, pages 4296–4297.

[Volkova and Yarowsky, 2014] Volkova, S. and Yarowsky, D. (2014). Improving gender prediction of social media users via weighted annotator rationales. In *NIPS 2014 Workshop on Personalization*.

[Waller, 2003] Waller, A. A. (2003). A dynamical systems model for investigating diversity. In *33rd Annual Frontiers in Education, 2003. FIE 2003.*, volume 2, pages F3D–1. IEEE.

[Wang et al., 2013] Wang, H., Li, Y., Feng, Z., and Feng, L. (2013). Retweeting analysis and prediction in microblogs: An epidemic inspired approach. *China Communications*, 10(3):13–24.

[Wang et al., 2016a] Wang, R., Liu, W., and Gao, S. (2016a). Hashtags and information virality in networked social movement: Examining hashtag co-occurrence patterns. *Online Information Review*, 40(7):850–866.

[Wang et al., 2016b] Wang, Y., Liu, J., Huang, Y., and Feng, X. (2016b). Using hashtag graph-based topic model to connect semantically-related words without co-occurrence in microblogs. *IEEE Transactions on Knowledge and Data Engineering*, 28(7):1919–1933.

[Wasko et al., 2005] Wasko, M. M., Faraj, S., et al. (2005). Why should i share? examining social capital and knowledge contribution in electronic networks of practice. *MIS quarterly*, 29(1):35–57.

[Wasko et al., 2004] Wasko, M. M., Faraj, S., and Teigland, R. (2004). Collective action and knowledge contribution in electronic networks of practice. *Journal of the Association for Information Systems*, 5(11):15.

[White, 2010] White, M. (2010). Clicktivism is ruining leftist activism. *The Guardian*, 12(August):2010.

[Wilcoxon, 1945] Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics bulletin*, 1(6):80–83.

[Williams et al., 2002] Williams, A., Turrell, P., and Wall, R. (2002). Let's twist: creating a conducive learning environment for women. *International Journal of Engineering Education*, 18(4):447–451.

[Xing et al., 2016] Xing, C., Wang, Y., Liu, J., Huang, Y., and Ma, W.-Y. (2016). Hashtag-based sub-event discovery using mutually generative lda in twitter. In *Thirtieth AAAI Conference on Artificial Intelligence*.

[Yadav et al., 2014] Yadav, D., Singh, R., Vatsa, M., and Noore, A. (2014). Recognizing age-separated face images: Humans and machines. *PloS one*, 9(12):e112234.

[Yang, 2016] Yang, G. (2016). Narrative agency in hashtag activism: The case of# blacklivesmatter. *Media and Communication*, 4(4):13.

[Yang et al., 2010] Yang, Z., Guo, J., Cai, K., Tang, J., Li, J., Zhang, L., and Su, Z. (2010). Understanding retweeting behaviors in social networks. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, pages 1633–1636. ACM.

[Young, 2018] Young, A. G. (2018). Using ict for social good: Cultural identity restoration through emancipatory pedagogy. *Information Systems Journal*, 28(2):340–358.

[Zaman et al., 2010] Zaman, T. R., Herbrich, R., Van Gael, J., and Stern, D. (2010). Predicting information spreading in twitter. In *Workshop on computational social science and the wisdom of crowds, nips*, volume 104, pages 17599–601. Citeseer.

[Zhang et al., 2015a] Zhang, J., Tang, J., Li, J., Liu, Y., and Xing, C. (2015a). Who influenced you? predicting retweet via social influence locality. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 9(3):25.

[Zhang et al., 2015b] Zhang, Q., Gong, Y., Guo, Y., and Huang, X. (2015b). Retweet behavior prediction using hierarchical dirichlet process. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.

[Zhang et al., 2016] Zhang, Q., Gong, Y., Wu, J., Huang, H., and Huang, X. (2016). Retweet prediction with attention-based deep neural network. In *Proceedings of the 25th ACM international on conference on information and knowledge management*, pages 75–84. ACM.

Curriculum Vitae

Education

- **Ph.D.** September 2012 - April 2020
  Information Technology (Computer Science Coursework)
  George Mason University, Fairfax, VA
  Advisor: Dr. Aditya Johri, Co-Advisor: Dr. Hemant Purohit
  Thesis Topic: *"Real-time Analytics for Resource Mobilization in Engineering Diversity Hashtag Campaigns"*
- **Master of Science** September 2005 - December 2008
  Computer Engineering, Artificial Intelligence and Robotics
  University of Tehran, Tehran, Iran
  Thesis Topic: *"Improving Reinforcement Learning Using Temporally-Extended Concepts"*
- **Bachelor of Science** September 2000 - September 2005
  Computer Engineering, Software Engineering
  University of Tehran, Tehran, Iran
  Project Topic: *"Document Retrieval Using Neural Network"*

Skills

Languages: Python, Java, C, SQL, Hadoop/Spark, Shell Script, Assembley, LaTeX
Statistics: R, Excel, Weka, Matlab
Applications: Intellij IDEA, Jupyter, PyCharm, Eclipse, , Rational Rose, Adobe Creative Suite
Operationg Systems: Windows XP / Vista / 7 / 8, Linux/Unix, Mac OS X

Related Work Experience

- **Graduate Research Assistant** **August 2016 - Present**

  Engineering Education and Cyberlearning (ECCL), George Mason UniversityFairfax, VA

- Analyzed unstructured/structured data from twitter (tweets/tweet meta data) for STEM-related issues
- Applied different machine learning methods to gain insight into the available dataset (*Gender Classification/Sentiment analysis*)
- Developed a machine learning framework for gender inference (individual (female, male), organization) on Twitter using tweet text (LIWC) and image (deep learning) with machine learning algorithms (SVM, Random Forest)
- Developed a machine learning framework for retweetability prediction on Twitter focused specifically on the content and graph structure of the hashtags in hashtag-driven campaigns
- Developed the topic and trend detection framework to compare the discussion themes in StackExchange vs Reddit online forums in social science subcommunities.
- Funded by NSF: EAGER: Social Media Participation as Indicator of Actors, Awareness, Attitudes, and Activities Related to STEM Education

- **Graduate Research Assistant** **September 2012 - May 2014**

  Center for Social Complexity (CSC), Krasnow Institute for Advanced Study Fairfax, VA

  - Worked closely with the team to implement in object-oriented design
  - Tested and debugged the previously written codes using *JUnit*
  - Managed research resources using *Zotero*
  - Wrote shell scripts to run simulations on computer cluster (ARGO)
  - Developed UI for end-users of the project using *SWING/Java*

  1. ONR/MURI Project: Developing new agent-based spatial simulation models for analyzing scenarios of societal consequences of disasters in the East Africa region

  2. NSF/CDI Project: Cyber-enabled understanding of complexity in socio-ecological systems via computational modeling

- **Software Engineer** **August 2011 - June 2012**

  Sharif University of Technology *Tehran, Iran*

  National Automation of Taxation System Project

  - Supervised the requirement analysis and design phases of the project
  - Executed system-wide testing scenarios to validate the software specification

- **Robot Programmer** **September 2006 - December 2006**

AI & Robotics Lab, University of Tehran *Tehran, Iran*

Dashboard design project

- Programmed a 6 DoF industrial arm robot, ABB-IRB 140, to model the dashboard of a vehicle

- **Requirement Analyst** **May 2002 - September 2002**

Mohaseb Software Company *Tehran, Iran*

Automation Software Project for the Deputy of Ministry of Economy in the Public Companies Affairs

- Extracted the desired behavior of the system by meeting domain experts
- Developed prototypes based on the requirements using *VB*

Additional Work Experience

- **Graduate Teaching Assistant** **January 2015 - August 2016**

Computing Resource, The Volgenau School of Engineering, George Mason UniversityFairfax, VA

- Provided on-site technical support regarding computing resources
- Troubleshot software/hardware issues for workstations in computer labs
- Monitored the quality of distance education lectures

- **Manager Assistant** **March 2009 - March 2011**

Tehran Municipality Information and Communication Technology OrganizationTehran, Iran

E-ticketing project for public transportation of the city of Tehran

- Developed the model of integrated electronic transaction
- Maintained the computer systems in bus depots in terms of hardware and software
- Designed the interface of two transportation systems, (i.e.: bus with metro)

- **Editor** **September 2005 - September 2006**

Ghalamchi Educational Institute *Tehran, Iran*

- Edited the preparatory questions for national M.Sc. computer engineering entrance exam

---

Honors & Awards

- Awarded Doctoral Fellowship for University of Tehran, Iran, 2005
- Ranked **Top 1%** (4th among more than 7,200) in the National University-Student Olympiad in Computer Engineering, 2005.
- Ranked **Top 1%** (603rd among more than 360,000) in the National University-Entrance Exam, 2000.

Teaching & Professional Experience

- **Lecturer** **September 2006 - June 2008**

  Iran University of Industries & Mines *Tehran, Iran*

  Courses taught: Computer network, System installation (Linux OS), Shell scripting in Linux

- **Lecturer** **September 2007 - June 2008**

  AmirKabir College of Management and Technology *Tehran, Iran*

  Courses taught: Computer laboratory (Linux operating system), Professional English

- **Reviewer** **October 2006**

  University of Tehran *Tehran, Iran*

  Reviewed the papers submitted for Student Conference on Electrical Engineering (SCEE)

Publications

Google Scholar

13. **Habib Karbasian**, Aditya Johri: "INSIGHTS FOR CURRICULUM DEVELOPMENT: IDENTIFYING EMERGING DATA SCIENCE TOPICS THROUGH ANALYSIS OF Q&A COMMUNITIES”, *The 51st ACM Technical Symposium on Computer Science Education*, USA, March 2020

12. Aqdas Malik, Yisheng Li, **Habib Karbasian**, Juho Hamari, Aditya Johri: ”LIVE, LOVE, JUUL: USER AND CONTENT ANALYSIS OF TWITTER POSTS ABOUT JUUL”, *Amarican Journal of Health Behavior*, USA, March 2019

11. Aditya Johri, Cassie Heyman-Schrum, Daniel Ruiz, Aqdas Malik, **Habib Karbasian**, Rajat Handa, Hemant Purohit: ”MORE THAN AN ENGINEER: INTERSECTIONAL SELF-EXPRESSIONS IN A HASHTAG ACTIVISM CAMPAIGN FOR ENGINEERING DIVERSITY”, *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, USA, 2018

10. Aqdas Malik, Aditya Johri, Rajat Handa, **Habib Karbasian**, Hemant Purohit: ”#ENGINEERSWEEK: BROADENING OUR UNDERSTANDING OF COMMUNITY ENGAGEMENT THROUGH ANALYSIS OF TWITTER USED DURING THE NATIONAL ENGINEERS WEEK”, *Proceedings of 125th ASEE Annual Conference*, USA, 2018

9. Aqdas Malik, Aditya Johri, Rajat Handa, **Habib Karbasian**, Hemant Purohit: ”#ILookLikeAnEngineer: USING SOCIAL MEDIA BASED HASHTAG ACTIVISM CAMPAIGNS AS A LENS TO BETTER UNDERSTAND ENGINEERING DIVERSITY ISSUES”, *The Collaborative Network for Engineering and Computing Diversity* (*CoNECD*), USA, 2018

8. Aqdas Malik, Aditya Johri, Rajat Handa, **Habib Karbasian**, Hemant Purohit: ”HOW SOCIAL MEDIA SUPPORTS HASHTAG ACTIVISM THROUGH MULTIVOCALITY: A CASE STUDY OF #ILOOKLIKEANENGINEER”, *First Monday Journal* USA, 2018

7. **Habib Karbasian**, Rajat Handa, Hemant Purohit, Aqdas Malik, Aditya Johri: ”REAL-TIME INFERENCE OF USER TYPES TO ASSIST WITH MORE INCLUSIVE SOCIAL MEDIA ACTIVISM CAMPAIGNS”, *AAAI/ACM Conference on AI, Ethics, and Society*, USA, 2018

6. Aditya Johri, **Habib Karbasian**, Aqdas Malik, Rajat Handa, Hemant Purohit: ”HOW DIVERSE USERS AND ACTIVITIES TRIGGER CONNECTIVE ACTION VIA SOCIAL MEDIA: LESSONS FROM THE TWITTER HASHTAG CAMPAIGN #*ILookLikeAnEngineer*”, *HICSS Conference*, USA, 2018

5. Alireza Saniani, **Habib Karbasian**: “DGPS AND ITS UTILITIES IN PUBLIC SERVICES”, *ECITY Conference*, Iran, 2009

4. **Habib Karbasian**, Majid N. Ahmadabadi, Babak N. Araabi: “CONCEPT EXTRACTION USING TEMPORAL-DIFFERENCE NETWORK”, *EUROCON*, Russia, 2009, pp. 1888-1894, DOI: 10.1109/EURCON.2009.5167904

3. **Habib Karbasian**, Majid N. Ahmadabadi, Babak N. Araabi: “IMPROVING REINFORCEMENT LEARNING USING TEMPORAL-DIFFERENCE NETWORK”, *EUROCON*, Russia, 2009, pp. 1716-1722, DOI: 10.1109/EURCON.2009.5167875

2. **Habib Karbasian**, Siavash Kayal: “DOCUMENT RETRIEVAL USING FUZZY MODELING OF NEURAL NETWORK”, *IADIS Conference*, Greece, 2008, pp. 363-367

1. **Habib Karbasian**, Parisa Rashidi: “PBT: PERSIAN PART OF SPEECH BRILL TAGGER”, *IADIS Conference*, Greece, 2008, pp. 348-352