Understanding Election Dynamics: A Graph Analysis Approach

Ronit Kumar Kataria – rk06451
Muhammad Azeem Haider – mh06858
Faraz Ali – fa06396

March 14, 2024

Instructor: Dr Qasim Pasta

Contents

1 Objective

The right to vote is a fundamental right of humans, however, the election process in Pakistan is nothing less than complicated. The relations between the different political groups and constituencies remain a mystery. While some claim that Pakistan as a country is getting politically aware, others refute the claim by citing the low turnover in voters. How true are the claims of hereditary seats in the national assembly?

The goal of this project is to leverage graph analytics and implement graph enhanced machine learning algorithms to learn more about how the election process in Pakistan has historically worked. Understanding the hidden complexities of elections in Pakistan for the past 30 years will also help in predicting the future results of elections in Pakistan. This project aims to highlight the connections between candidates, parties, and voters by modeling election data as a graph. The graph representation of an otherwise common relational data will help us uncover relations, clusters, communities, and truths that otherwise are not apparent in a relational database.

2 Methodology

For our project we will be using data from the general election of Pakistan in 1993 to election data in 2023. The following will be the nodes in our project:

- Candidate
- Constituency

The relationships/edges in our graphs will be **contested** from candidate name to constituency number, and **party members** like the co-actor network.

The attributes such as registered voters of the constituency numbers will help us analyze the trends in the number of actual voters voting compared to registered voters helping us better understand if Pakistan as a country are becoming more politically aware. We will also be able to predict the number of future voters through this metric. Furthermore, having prior data of individuals of specific political party winning from a constituency will help us predict which political party will prevail in the next elections.

As general analytics, it would also be interesting to see if particular political parties make clusters/communities around a few constituencies, thus highlighting their dominance in the area. It would also be interesting to see which **constituency** node becomes isolated if we only include the top 3 biggest political parties of Pakistan. This would highlight what constituencies in Pakistan are usually dominated by individual candidates not associated to any political party. Additionally, detection of patterns in the graph network, such as recurring relationships between specific candidates and constituencies across multiple elections, will unveil long-standing political dynamics or trends.

The data is available to us from election in 1993 to 2018. We will be collecting data for the election in 2023, and build the graph database using the csv.

3 Potential Outcomes

- **Influential Political Leader** With graph database, we can easily identify influential candidates, parties and constituencies within system
- **Community Detection** We also aim to classify different types of voters, based on their voting dynamics so we can separate out their clusters to bring more informative outcome.
- **Predictive Modelling** With extensive amount of data, We want to make models that can predict election results based on historical trend and also figure out how different things affect those results.
- **Election System** We also aim to explore the political analysis that lead to distribution of political power and support which help policymakers to draw conclusion across constituencies.

4 Expected Challenges

When dealing with the complexities of election data, there are problems that must be overcome, such as filtering out noise and maintaining complicated interactions within the graph. Our goal is to make sure that the data is accurate by cleaning it up and concentrating on the details that are essential. With data up to 2018, our objective is to acquire data from 2023 in order to gain a more in-depth understanding of the electoral dynamics in Pakistan.