Project Specification

Project title:	Project Title here
Student:	Tristan Thomson 80720305
Supervisor:	Dr Ida Pu

Work overview

The European Parliament holds a key role in the legislative process of the European Union. It discusses, negotiates and votes on legislation put forwards by the European Commission. The parliament is composed of 751 members (MEPs) who represent the 28 member states of the European Union.

I plan to analyse the influence of Members of the European Parliament's (MEP) social networks on the outcome of their votes on a motion in the European Parliament. I would like to infer "unknown influences" on these MEPs which cause them to vote in unexpected ways and potentially visualise the results of my analyses. The project could ultimately predict future votes by matching shared characteristics between MEPs.

I am thinking of analysing the outcome of the Directive on Copyright in the Digital Single Market from a publicly available spreadsheet of vote results. (Existing dataset includes: First & Last names, votes, European-level political groups, National-level political groups, Jurisdiction, Country of origin, EP webpage, twitter handle (if applicable))

I am hoping to achieve a series of measures which will hopefully indicate why particular MEPs vote in ways which were not otherwise predictable, leading to the possibility of predicting future votes from combinations of those characteristics. Completing this project will necessitate a combination of social network analysis, graph theory and probability.

Technologies and Materials

I plan on using Python for most of the work. It seems a particularly good choice for its wide range of supporting libraries and excellent documentation. I will source data from European Parliament websites with pre-constructed datasets as well as collecting data myself by using social media APIs. I intend to construct and analyse the data using Pandas and NumPy, construct social networks using NetworkX, a Python package for graph theory and predict future votes with Machine Learning techniques using TensorFlow.

- Tweepy Twitter API library : here
- Tensorflow : here
- NetworkX : here
- Pandas : here
- NumPy : here

Motivation research

My personal interest in politics and the systems behind passing legislation along with my interest in Data Science have led me to choose this project. I would like to expand my knowledge of the European Parliament as well as to work on a technically challenging project.

The other aim is to explore the notion of transparency in the European Parliament. I believe transparency is central to a well-functioning legislative process and while measures have been taken to record information about certain influences acting on European representatives, there is still space for improvement. We, as European citizens have chosen our representatives because they say we share the same point of view, if they vote otherwise, the absence of transparency is hurtful. I would like to use my skills to assist this cause.

0.1 Motivational Resources(links)

Social Networks in Policy Making

U.S. Senators' Social Network Analysis Based on Twitter

European Transparency Register

#SaveYourInternet fight against Article 17 [ex Art. 13] continues

Directive on Copyright in the Digital Single Market vote results

Existing knowledge

I will draw on Python experience from my 1st year module Problem solving for Computer Science. I will make use of my past studies in 2nd year Algorithms and Data Structures as well as my current 3rd year module Advanced Algorithms and Data Structures for the graph theory and probability in my project. I will also rely on my 3rd year Neural Networks, Machine Learning and Data Mining modules for the completion of this project.

New knowledge

I will need to gain knowledge of data collection, particularly extracting data from social media APIs and rendering it useful through manipulation. This project will also require me to learn about Social Network Analysis, Graph Theory, Probability and Machine Learning. I will acquire some of these skills from my current modules but will likely need to explore the others more in depth using mainly resources found online.

Timeline and milestones

My project will be stored on Github at : https://github.com/TristanThomson/Year-3-Final-Project

- MVP: Data collection, establishing social influence networks around "key rebel" MEPs, defining metrics to use for weighting the edges between individuals Data Visualisation.
- Backlog 1: Using probability to infer other influences acting on the MEPs' voting.
- Backlog 2: Prediction based on known data using Machine Learning techniques.

Week Monday	**Topic**	**Milestone**
30/09/2019		Ideation Submission
07/10/2019	Planning & research	
14/10/2019	Planning & research	
21/10/2019	Planning & research	
28/10/2019	Concept validation	Project Specification Submission
04/11/2019	Data Collection	
11/11/2019	Data Collection	
18/11/2019	Finding key rebel MEPs	
25/11/2019	Metrics research for SNA	
02/12/2019		
09/12/2019		
16/12/2019		
23/12/2019		
30/12/2019		
06/01/2020		
13/01/2020		Interim Project Report
20/01/2020		
27/01/2020		
03/02/2020		
10/02/2020		
17/02/2020		
24/02/2020		
02/03/2020		
09/03/2020		
16/03/2020		
23/03/2020		Draft report submission
30/03/2020		
06/04/2020		
13/04/2020		
20/04/2020		
27/04/2020		
04/05/2020		
11/05/2020		
18/05/2020		Final report submission