Department of Political Science

Analysing Public Sentiment and Hate Speech on Reddit: Understanding the Impact of Immigration in Ireland

Author: Shekhar Kedia

Supervisor: Tom Paskhalis

Dedicated to all those who embrace diversity

and strive to create a more inclusive

and empathetic society for all

School of Social Sciences and Philosophy
Assignment Submission Form

Student Name:	**Shekhar Kedia**
Student ID Number:	**23351315**
Programme Title:	**MSc. Applied Social Data Science**
Assessment Title:	**Dissertation**
Lecturer (s):	**Tom Paskhalis**
Date Submitted	**12th Aug 2024**

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at: http://www.tcd.ie/calendar

I have also completed the Online Tutorial on avoiding plagiarism 'Ready, Steady, Write', located at http://tcd-ie.libguides.com/plagiarism/ready-steady-write

Signed: 

Date: **12th Aug 2024**

Acknowledgement

I would like to express my heartfelt thanks to Trinity College Dublin for providing the essential resources and support throughout this study. My deepest appreciation goes to my supervisor, Tom Paskhalis, whose expert guidance and encouragement were invaluable in completing this work. I am also profoundly grateful to my friends and family, especially my parents, for their unconditional love and support. Lastly, I extend my gratitude to the human annotators for their meticulous work in developing the hate speech classifier and to the Reddit community for their contributions.

Abstract

This study explores the association between evolving online sentiment and hate speech towards immigrants in Ireland, particularly in the context of increasing refugee arrivals in the country. To examine the link between rising immigration in Ireland and the spread of hate speech on Irish subreddits, this study adopts a three-tier analysis framework. First, it compares mentions of immigrants in Reddit posts and comments from 2010 to 2024 with immigration trends. Second, it measures the net sentiment (positive minus negative) of these posts and comments using a sentiment dictionary. Third, using supervised machine learning it classifies posts related to immigrants as hate speech or not, and assess its prevalence. Topic modeling is further explored to discover common themes in posts identified as hate speech in relation to immigrants.

The results show an increased in mention of immigration in Reddit posts and comments during the study reference period. It also shows a growing negative sentiment (using the sentiment dictionary) in online discourse, with a sharper decline in net sentiment value observed in Reddit comments post-2020. Further, while there remains scope for further refinement, the Support Vector Machine (SVM) model outperforms the Naïve Bayes in detecting hate speech in Reddit posts with an accuracy of 76%. The topic modeling method identifies some of the themes that are in tune with the current scenario in Ireland in context of immigration like “seek housing”, “travel visa and immigration enquiry”, “housing crisis and government role” and “Ukrainians refugees and support services provided.”

Keywords: Immigration, Refugees, Sentiment Analysis, Hate Speech, Social Media, Reddit, Machine Learning, Ireland, Support Vector Machine (SVM), Naïve Bayes, Topic Modeling

N.B.: The replication files are available here: https://github.com/shekhar-kedia/Dissertation-project_ASDS.git

Table of content

List of figures

List of tables

1 Introduction:

Ireland has experienced a significant surge of immigration over recent years. The Central Statistics Office, defines immigration *as the action by which a person establishes his or her usual residence in the State for a period that is, or is expected to be, of at least 12 months, having previously been usually resident in another Member State or a third country*[1]. The influx of immigration includes both state-dependent immigrants such as refugees and asylum seekers, and immigrants that contribute to the economy of the country as entrepreneurs and employees, essentially active in the labour market. This rise in population of state-dependent immigrants in the country is driven by multitude of factors like Russia-Ukraine conflict and rise of Taliban in Afghanistan while the increase in multinational corporations and shortage of local skilled labour force has led to rise in immigrant workforce. As a result, Ireland has seen an increasingly diverse population, with various nationalities and ethnicities arriving, particularly in the last decade.

This demographic shift has resulted a mixed response from local communities, creating both opportunities and challenges that influence public sentiment. Recent incidence of riots and burning of hotels dedicated to refugees and asylum seekers reflect growing tensions and wariness among the local population regarding immigration. Moreover, social media platforms have become a breeding ground for anti-immigration sentiments, where individuals often express extreme views, including racist remarks, threats, and harassment.

Moreover, social media platforms provide space for anti-immigration nationalists to spread hate content. On these platforms, individuals often express their opinions, emotions, and grievances, sometimes resorting to extreme and harmful tone. This has led to an environment where hate speech, including racist remarks, threats, and harassment, is increasingly prevalent. As social media platforms like Reddit become key spaces for expressing opinions, understanding the nature of these online discussions is crucial. This research explores the relationship between immigration and public sentiment in Ireland, with a focus on identifying and analyzing hate speech through advanced machine learning techniques.

The study aims to explore how rising immigration trends have impacted public sentiment, particularly in the context of hate speech expressed on Reddit over the past 15 years. By employing advanced computational methods, including supervised machine learning, this research seeks to detect, measure, and analyse hate speech, offering insights into the broader implications of immigration on Irish society.

It is firstly important to understand the immigration landscape in Ireland and how public sentiment has evolved in the past 10-15 years. Besides, through extensive literature review, how racial expressions and hate speech especially in the context of immigration is spread online and its implications will be explored.

Immigration landscape in Ireland:

The 2022 Census shows that Ireland is a highly diverse country with one in every five residents born outside (migrant) the country (Central Statistics Office, 2023). Again, at least three out of five migrants originate from the UK or other EU countries (McGinnity, 2023). It is interesting to note that the migrants from the UK and EU don't require any form of permission to live and work in Ireland while, the non-EU migrants need special authorisation and valid work permit. The data from Eurostat 2023 shows that the non-EU migrants primarily move to the country for work, high-skilled jobs and educational purposes which explains why the migrant group have a higher education and income-levels than the average Irish population (McGinnity, 2023). Notably, in recent years, there has been a significant increase in the non-EU labor force to address labor shortages in Ireland (Murphy & Sheridan, 2023).

The immigration policies of the Irish government have also made the country a favourable destination to people seeking protection such as refugees and asylum-seekers. Understanding key terminologies related to migration in Ireland is crucial. The three main categories of migrants in the Irish context are asylum seekers, refugees, and migrant workers. The Citizens Information portal defines asylum seeker as a person who has left their country and is seeking protection in Ireland (Citizens Information, 2024). They are also called international protection applicants where the Irish state provides accommodation, food and medical care while the decision

[1] Background Notes Population and Migration Estimates, April 2023 - Central Statistics Office

on the application is under consideration. Refugee status is granted to a person after the application has been processed and it is confirmed that there is fear of persecution in their own country because of their race, religion, nationality, political opinion or membership of a particular social group. In case of Ukrainian refugees, there is a separate directive called temporary protection which is an EU law introduced in 2001. The directive has a special procedure to deal with a 'mass influx' of people that are in need of international protection. Under this directive, without the need for individual applications for international protection, the EU countries grant immediate protection to people whereby they are provided residence permit for a minimum of one year and a maximum of three years besides getting access to employment, accommodation, social welfare, healthcare, and education. Notably, the directive was triggered by the EU commission for the first time in March 2022 in response to the war in Ukraine (Official Journal of the European Union, 2022).

On the other hand, National Youth Council of Ireland (NYCI) defines migrant worker as someone who is working in a state of which s/he is not a national and requires a valid work permit to be employed in a specific job (National Youth Council of Ireland, 2024). In addition to this, young people who come to the country for education are issued student visa (Stamp 2 and 2A) and are registered with the Garda National Immigration Bureau (GNIB). It is also important to note that the study focuses only on legal migrants i.e., someone who has a documented migration status.

The year 2022 has been a critical year for Ireland where the number of people seeking protection is by far the highest in Irish history. It is also a period in which the Irish economy recovered from the COVID-19 pandemic leading to increase in non-EU permits (Murphy & Sheridan, 2023). The Central Statistics Office estimated the total number of immigrants entering the country around 107,800 in the year to April 2022, while the number of people leaving the country i.e., emigrating was estimated to be 59,600, creating a positive net migration of 48,200. The table below gives an indication of new immigrants coming into the country since 2010 taken from the Central Statistics Office (CSO) portal[2]. It is important to note that these are estimated figures. The table also highlights the number of refugees and asylum seekers provided by The UN Refugee Agency[3].

Table 1: Estimated number of new immigrants entering Ireland (including refugees and asylum seekers) for the period 2010-2024

Year	Total estimated immigrants (in April)	Total refugees	% of refugees of total immigrants	Total asylum seekers
2010	41,800	9,099	22%	5,088
2011	53,300	8,234	15%	5,380
2012	57,300	6,309	11%	5,423
2013	62,700	5,991	10%	5,464
2014	66,500	5,832	9%	4,576
2015	75,900	6,108	8%	4,982
2016	82,300	5,720	7%	4,221
2017	95,300	6,394	7%	5,991
2018	96,000	6,012	6%	7,159
2019	97,100	7,795	8%	7,853
2020	95,600	9,035	9%	7,393
2021	74,100	9,571	13%	6,904
2022	107,800	81,256	75%	15,108
2023	141,600	113,902	80%	22,250

[2] https://data.cso.ie/table/PEA24
[3] unhcr.org/refugee-statistics/download/?url=FU89Cy

Ireland's response to the Ukraine crisis included significant support and services from the Department of Social Protection (DSP), with Ukrainian refugees being distributed across various counties The CSO maintains a separate database to track the location of Ukrainian refugees across the country[4]. The map shows the spread of the Ukrainian refugees in the country as taken from the CSO website. It is to be noted that in 2022, of the total number of refugees coming to the country, 79% are Ukrainians and in 2023, it is 80%. This goes to show that most of the recent refugees in the country are Ukrainians.

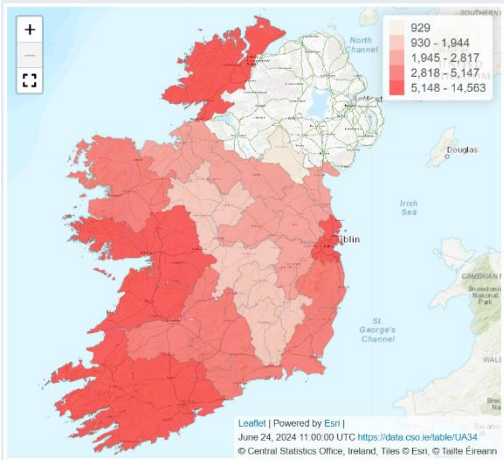

Figure 1: Map highlighting the county wise spread of Ukrainians in the country

Ireland has faced a long-standing housing crisis and with the influx of large number of immigrants, it is but obvious the country experienced a serious shortage of affordable housing besides increasing cost of living and inflation issues. The recent report by The Housing Commission highlights some of the critical issues the country's housing sector faces (The Housing Commission, 2024). The Social Justice Ireland, an independent think tank body, that focuses on issue of social justice and public policy states that there is a substantial gap between population growth in the country and housing availability leading to a situation where for each new housing unit, on average there are at least four additional individuals highlighting significant issues in the supply of housing (Addressing Ireland's Housing Crisis: Urgent Policy Reforms Needed, 2024).

The housing crisis has resulted in a steep increase in rental and housing prices, particularly in urban areas. This has led to large number of young adults (aged 25-34) living with their parents (Disch & Slaymaker, 2023). The crisis has also led to rise in homelessness situation. According to Peter McVerry Trust, a national housing and homeless charity organization, as of March 2024, 13,866 people are homeless i.e., accessing emergency accommodation in the country (Peter McVerry Trust, 2024).

Public attitudes on rising immigration and the role of social media in expressing related concerns:

Amidst Ireland's housing struggles, there is growing concern among the public that asylum seekers and refugees are being prioritised over native citizens (Kumar & Donoghue, 2023). Besides, putting strain on housing market, immigrants are also seen as a threat to services and while Ireland lacks a significant anti-immigrant political movement or a far-right political party, there have been recent incidences which suggest hostile attitudes of general population towards immigrants. There have been many protests against accommodation centres being allotted to asylum seekers in recent years (Keena, 2022) (Deegan, 2023) (Slater & Lucey, 2023) including arson attacks on hotels that were supposed to accommodate asylum seekers (Lally & J, 'Fire at Dublin building intended as , 2024) (Lally & Fallon, 2023). Furthermore, there was a riot in Dublin city centre on 23rd Nov 2023 that was allegedly led by the far-right ideologists as a reaction to stabbing of a woman and three children by a man. While these incidences cannot be ascertained to anti-immigrant attitude of general Irish population, it is a reflection of growing hatred and hostility towards immigrants.

Social media platforms have become critical infrastructures for the production and dissemination of racism, particularly by anti-immigrant and racist individuals (Matamoros-Fernández, 2017). Sharma (2018) discusses the prevalence of "ambient racism" on social media which is not different from the explicit form of racism that is known to all. He highlights that this type of racism occurs at a micro-level in user comments, posts, and interactions, forming part of everyday internet communication. Radical right people (read anti-immigrants) that are part of online groups play a significant role in shaping public opinions on immigrants and refugees. However, the 'dynamics of these online groups remain poorly understood' (Törnberg & Wahlström, 2018). The problem is the content shared by these group of people feeds into other users' emotions of insecurity and further generate anti-immigrant attitude among the masses. This strategy is effective in spreading anti-immigrant discourse and normalizing racial slurs and racist expressions when ordinary users interact and communicate regularly with

[4] https://data.cso.ie/table/UA34

far-right and racist people (Ekman M. , 2019). It is not wrong to say that online social media platforms like Facebook and Reddit act as medium to garner xenophobic, nationalistic, and racist discourse which results in shaping negative sentiment and attitudes towards immigrants (Farkas, Schou, & Neumayer, 2018). When other users of these social media platforms start interacting with the anti-immigrant far right group of people, the xenophobic and racist attitude among the users tend to amplify (Ekman M. , 2019). In December 2017, a woman from Sweden was convicted of online hate speech when she posted abusive comments about Somali immigrants on Facebook. It was found that she was 'brainwashed' and motivated by the hateful content shared by other users in the 'Stand up for Sweden' (SufS) group (Larsson, 2018).

When racial expressions become mundane and mixed with other contents, it tends to change the construct and perception of immigration, leading to increased hostility towards them. Miles and Brown (2003) emphasise that a recurrent element is to depict violence against women as an 'immigrant problem'. Recent forced migration into Europe provides another example of altered image of immigrants and leading to anti-immigration sentiment where male refugees are seen as sexists and rapists (Ekman M. , 2018). The infrastructure of these social media platforms also fails to deregulate hate speech towards immigrants especially because of vague user policies and the fact that moderation is often outsourced and decontextualised (Nikunen, 2018).

The literature makes it clear that hate speech and racist expressions on social media do not exist in vacuum and are manifested by people's actions when they partake in the production and circulation of such content (Ekman M. , 2019). Anti-immigrant and racists people take advantage of social media platforms to normalise usage of their radical right-wing sentiments which is filled with hate towards immigrants (Wahlström & Törnberg, 2019).

1.1 Contribution and study objectives:

The literature review throws light on the complex immigration situation in Ireland and how public sentiment is affected because of it. It is seen that with rise in numbers of new immigrants coming to the country, people are more wary of immigrants. The literature also highlights how anti-immigrant and racist content is produced and spread on mainstream social media platforms and normalising use of racial slurs and expressions.

However, the existing literature does not sufficiently explore the relationship between the recent rise in immigration numbers in Ireland and the spread of hateful and racist content on Irish social media platforms. This study is a novel attempt in understanding this relation and also exploring if there is any differential effect because of rise in state dependent immigrants. Besides, the study will focus on highlighting county-level findings as it is seen that there is a variation in spread of immigrants, especially refugees across the 26 counties in the country.

The key research question that the study will aim to answer is: How has the rise in immigration to Ireland over the past 15 years influenced public sentiment, and to what extent is this reflected in the prevalence and nature of hate speech on social media platforms like Reddit?

The primary study objectives stemming from the research question that will be explored are:

- Measure the number of mentions of immigration and immigrants in Reddit posts and comments during the reference period and analyse the association with rising immigration, including county-level changes in net migration and the distribution of refugees.
- Assess the net sentiment of user-generated Reddit content (posts and comments) related to immigrants during the reference period and examine its association with rising immigration, including county-level changes in net migration and the distribution of refugees.
- Develop a machine learning model to classify content related to immigrants as hate speech or not and measure the prevalence of hate speech during the reference period. The study will also explore the association between hate speech prevalence and rising immigration, including county-level changes in net migration and the distribution of refugees.

2 Approach and methodology:

To explore the relationship between rise in immigration in Ireland and the spread of hateful and racist content on Irish subreddits the study will adopt a three-tier framework for analysis. Firstly, assessment and comparison of the number of mentions of immigrants in Reddit posts and comments over the last 15 years (2010-2024) with trends in immigration during that period will be done. Secondly, net sentiment (positive minus negative) of Reddit posts and comments related to immigrants will be measured using an imported dictionary. Lastly, using a supervised machine learning approach, classification of Reddit posts related to immigrants as hate speech or not will be done, determining the prevalence of hate speech during the reference period. Additionally, the study will use topic modeling to explore common themes in posts categorised as hate speech.

2.1 Scope of the study and data used:

This study utilises several datasets to explore the relationship between immigration and spread of hateful content on Irish social media. The primary data source consists of Reddit posts and comments from the official subreddits of all 26 counties in the Republic of Ireland, covering the period from January 2010 to May 30, 2024. Data was scraped[5] from each county's subreddit, with details on the specific subreddits provided in the Annexure 1. The original dataset scraped was in JSONL (JavaScript Object Notation Lines) file format that consists of multiple JSON objects separated by a newline. This was then processed to CSV (Comma Separated Value) files for ease of analysis. The scripts used for data processing are included in the replication folder. Using the Quanteda[6] package in R, the text was pre-processed through several steps: cleaning punctuation, removing numbers, hyphens, and blank spaces, and converting all text to lowercase for uniformity. Collocations were created as needed to capture contextually significant phrases. This process resulted in the creation of two document-feature matrices (DFMs): one for posts and one for comments, corresponding to all 26 counties.

The final dataset comprised of 107,207 posts and 1,286,835 comments. It is important to note that many posts and comments are deleted over time, either by Reddit moderators due to policy violations or by users themselves. The final counts reflect only the posts and comments that remained after such deletions.

In addition to the Reddit data, immigration statistics for the reference period were sourced from the Central Statistics Office (CSO). For county level analysis, direct records of immigrant numbers at the county level were unavailable, and so the study uses net migration data (number of migrants per 1,000 population) as a proxy as this is a good metric to understand the net increase in new people coming to the country over the reference period. The figures are mentioned in Annexure 2 for reference.

To draw findings on state-dependent immigrants and discussion on Reddit, the Integration Minister Roderic O'Gorman recently shared information of about 74,000 refugees and International Protection Applicants that are in State-funded accommodation (Irish Mirror, 2023). This is a good metric to understand the spread of state-dependent immigrants in various counties of the country. The figures are mentioned in the Annexure 3 for reference. Moreover, as 80% of the recent refugees in the country are Ukrainians, the study will also use the number of Ukrainians entering the country at the county level (refer Annexure 4) maintained by the CSO as an additional analysis.

[5] Download subreddit or user data (photon-reddit.com)
[6] Quantitative Analysis of Textual Data • quanteda

2.2 Methods used in the study:

2.2.1 Text analysis using Quanteda and Document-Feature Matrix (DFM):

The Quanteda-package (Benoit K. K., 2018) provides various methods for the quantitative analysis of textual data by creating a corpus. A bag of words approach (Jurafsky & Martin, 2008) is used where the first step includes breaking of all the words into tokens or individual words. It is then followed by cleaning the text and creating a document-feature matrix (DFM) to conduct statistical analysis, such as frequency analysis, and sentiment analysis using an imported dictionary. Each method as outlined by Benoit (2019) in a typical quantitative text analysis project was followed. A detailed process is present in the Annexure 5 for reference.

2.2.2 Dictionary method for sentiment analysis:

The dictionary method for sentiment analysis involves comparing texts with a predefined dictionary to determine positive and negative scores. The net sentiment score is calculated by subtracting the negative score from the positive score. According to Grimmer & Stewart (2013), “Dictionaries use the rate at which keywords appear in a text to classify documents into categories or to measure the extent to which documents belong to particular categories.” This method goes beyond merely counting word frequencies by associating tokens with predefined meanings.

For the purpose of the study, the Lexicoder Sentiment Dictionary (Young & Stuart, 2012) was used. It consists of a total of 4,567 positive and negative words. A sentiment dictionary comprises canonical terms or concepts ("keys") associated with a list of equivalent terms ("values"). The Lexicoder Sentiment Dictionary, originally aimed at classifying news content into positive and negative sentiment, is well-suited for sentiment analysis in this context as it employs several strategies for disambiguation like:

- Key Word in Context (KWIC): Determines context (e.g., “lie” is negative in the context of “a lie” or “lie to”)
- Polysemy Handling: Accounts for words with multiple meanings through contextual analysis or word-sense disambiguation (e.g., “my arm hurts” versus “the US sold arms to the Saudis”)
- Negative Phrase Detection: Identifies phrases like “not good” as negative

2.2.3 Supervised method for text classification:

Supervised machine learning involves using algorithms to classify items (documents) into predefined labels based on human-annotated training data (Sarker, et al., 2020). These algorithms learn patterns from labelled data and provide a model for predicting unseen data (Han, Pei, & Kamber, 2011). In the context of the study where Reddit content (posts) is to be classified as hate speech or not, a binary-class supervised classification method was used. The goal of the supervised learning is to train a machine model that maximises generalisation and minimises overfitting.

Components of supervised learning:

Supervised learning in this study involved several key components. First, a human-annotated dataset was divided into two subsets: a training set used to train the machine learning classifier, and a test set used to validate its performance. Various classifiers were employed to predict the labels of unseen data, including Naïve Bayes and Support Vector Machines (SVM). To assess the performance of these classifiers, validation methods such as the held-out method or cross-validation were used. Performance was then measured using a range of metrics, including the confusion matrix, accuracy, precision, recall and F1 score.

A detailed guide on how the training set was developed is provided in the Annexure 6.

Finalising and running the model:

The process of training and finalising a supervised learning model began with selecting appropriate classifiers, with this study focusing on Naïve Bayes and Support Vector Machines (SVM) due to their established effectiveness in handling text data. The next step involved validating the model to ensure it performed well on unseen data, avoiding overfitting. This was achieved through out-of-sample performance evaluation and k-fold cross-validation, where the model was trained and tested on different folds of the data.

Next tuning of parameters was done using techniques such as grid search and random search to identify the optimal parameter combinations. The model's out-of-sample performance was calculated for each combination, and the best-performing parameters were selected. Once the optimal parameters were determined, the final model was re-estimated using the entire training dataset. This finalised model was then used to generate predictions or labels of hate speech for new data.

The final step involved running the model to classify new data. The predictions were merged with the training data which was later used for further analysis.

2.2.4 Topic modeling:

Topic modeling represents a family of unsupervised classification algorithms designed to uncover the latent thematic structure within a large corpus of text. These models help in data reduction process by identifying key topics or themes that best represent the data's structure. The primary objective of topic modeling in this study is to understand the underlying themes present in the corpus. The corpus for this exercise will be all the Reddit posts that are classified as hate speech. This will help discover some of the key topics around which there is usage of hate speech related to immigrants. This method involves transitioning from observable words to latent topics, thereby providing a deeper and more meaningful understanding of the textual data.

For the purpose of the study Latent Dirichlet Allocation (LDA) method will be adopted which is a generative probabilistic model that assumes documents are mixtures of latent topics, with each topic represented by a distribution over words (M. Blei, Ng, & Jordan, 2001). Detailed steps to using LDA for topic modeling is provided in Annexure 7.

2.2.5 Statistical analysis:

The study also uses a mix of descriptive and inferential statistics for analysis. While broad trends over time are seen, correlation and regression analysis are done to understand the direction and strength of the association between variables. The key dependent variables used in the regression analysis are:

- Percentage of mentions about immigrants in Reddit posts and comments
- Net sentiment score of comments and posts
- Percentage of hate speech in Reddit posts that are related to immigrants

2.2.6 Tools used for analysis:

Rstudio using R language was used for statistical analysis, sentiment analysis and machine learning. Packages like *dplyr, tidyverse*, were used for data manipulation; *lm* for regression analysis; *stringi, stringr,* and *quanteda* for text processing; *caret, MLmetrics* for machine learning; *topicmodels* for topic modeling; *ggplot2* for visualisation.

2.2.7 Ethical considerations:

This study involved scraping data from Reddit, including users' comments and posts, to analyse public sentiment and the presence of hate speech related to immigration. Given the nature of the data, stringent ethical considerations were applied to ensure the privacy and anonymity of the individuals whose data was analysed. First and foremost, all data was anonymised before any analysis was conducted. Identifiers such as usernames, URLs, and other personal identifiers were carefully removed to prevent any potential identification of the users. The data was processed in a manner that ensured individual privacy was maintained at all times.

Moreover, all the analysis were conducted at an aggregated level rather than focusing on individual users. This approach helped respect user privacy and focus on broader trends rather than specific cases or individuals. The insights derived from the study pertain to collective sentiment and patterns of hate speech rather than any single user's opinions or behaviours.

3 Results:

3.1 Trends from Reddit data (2010-2024):

With the processed corpus of Reddit posts and comments during the study reference period, statistical visualisations of the data, such as histograms depicting the distribution of posts and comments over time, was generated. The graph below shows that the data distribution is negatively skewed, with a noticeable increase in activity after 2020. To address this imbalance in spread, the analysis was standardised on an annual basis. For example, the percentage of mentions of immigration in each year was calculated and compared across the years.

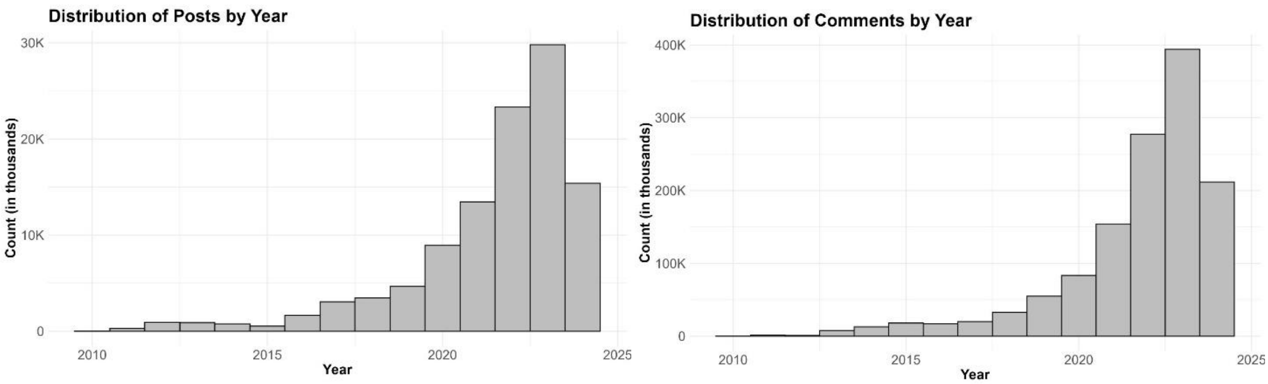

Figure 2: Distribution of Reddit Posts and Comments for the reference period

Doing a deep dive in text analysis to understand the data, ngrams for frequency of occurrence of words (otherwise known as features) over the reference period was done. The distribution graph below shows the top feature in the posts is “work”, followed by “good” and “time”. The word cloud gives a more vivid idea on the top words present in all the posts. It is observed that a lot of posts are around seeking “help”, “wondering”, “suggestions” or even “advice” which makes sense because a lot of people would put up questions to get peoples advice, opinions and suggestions on things.

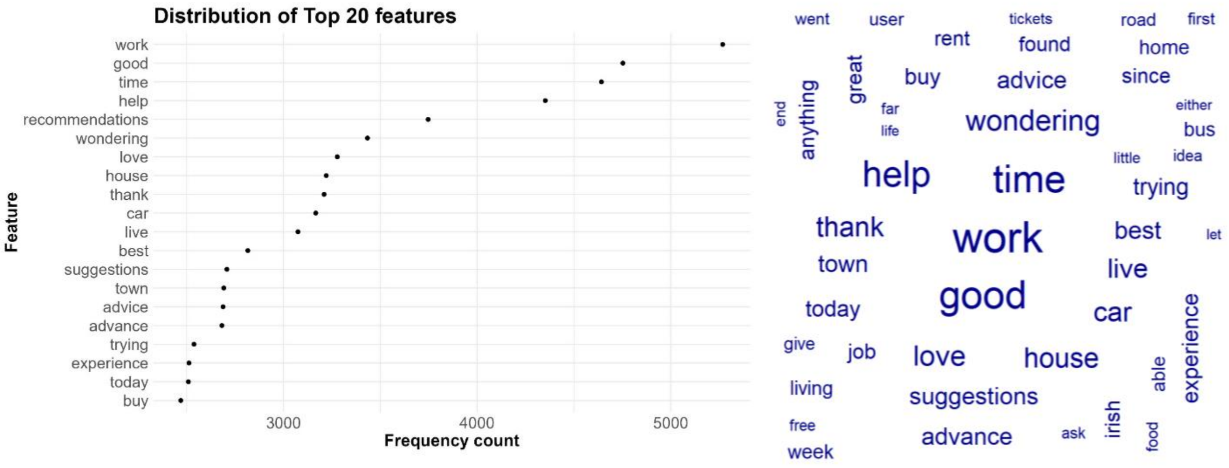

Figure 3: Distribution and word cloud of top features in Reddit Posts for the reference period

Similarly, when looking at the top feature list for the Reddit comments during the reference period, it is found that "time", "good", and "work" are still the most mentioned words in the Irish ecosystem on Reddit. Besides, a lot of features related to housing like "house", "home", "rent", "live" and "living" are also present. This is not out of the line as housing is a big topic of discussion in Ireland given the housing crisis.

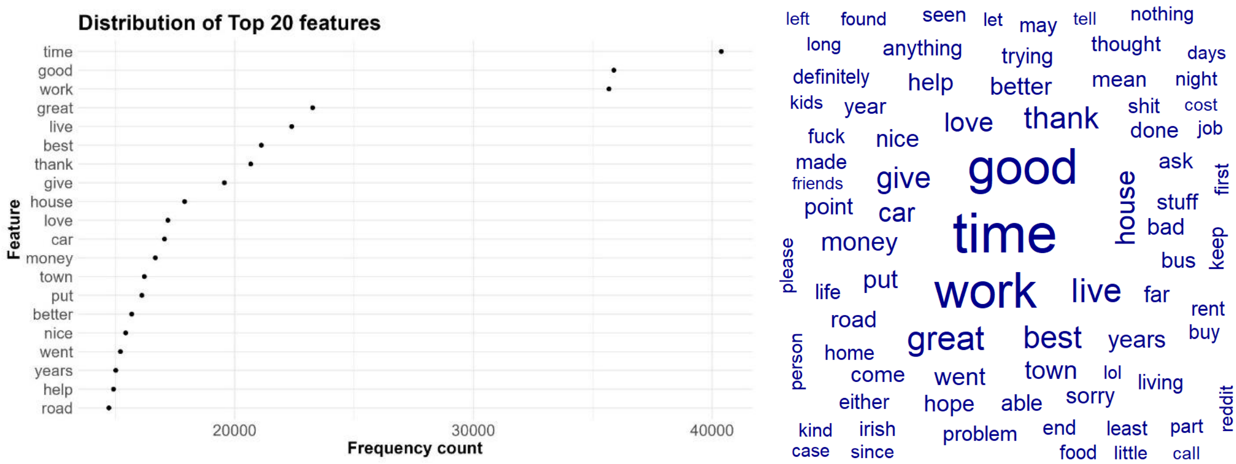

Figure 4: Distribution and word cloud of top features in Reddit comments for the reference period

Similarly, in the comments, words like "time," "good," and "work" are also prevalent, along with terms related to housing such as "house," "home," "rent," and "live." This aligns with Ireland's ongoing housing crisis, a major topic of conversation in the country.

3.1.1 Establishing credibility to the dataset:

To ensure the dataset's credibility, it is crucial to verify that the data accurately reflects real-world events. Given that this study aims to analyse public sentiment on social media regarding immigration, the data must align with significant global events. A benchmark for this is the COVID-19 pandemic, during which discussions about the virus were expected to spike, especially post-2020.

The chart below shows the pattern for posts and comments having mentions of 'COVID-19'. It can be seen that these discussions spiked 2020 onwards. Before 2020, less than 0.5% of the posts and comments had mentioned about "virus", "covid19", "vaccine", "corona", or "covid" and in 2020, the percentage of posts and comments having mentions about the pandemic jumped to more than four percent levels and two percent levels respectively. Post 2020, there is a gradual dip in percentage of mentions which is expected as the situation normalises. This trend suggests that the dataset is in tune with real-world events, establishing its credibility for further analysis.

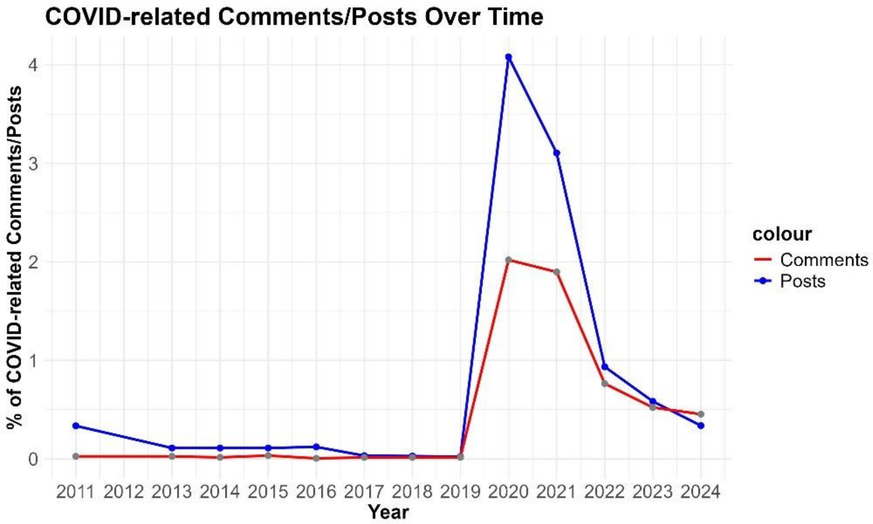

Figure 5: Trend of COVID-related mentions in comments and posts for reference period

3.2 Discussions and mentions around immigration:

As seen before, the data is in line with the real-world events and as expected, given the rise in number of immigrants in Ireland, the percentage of posts and comments having mentions of immigrants has gradually increased over the reference period. The lists the keywords used to filter posts and comments related to immigration in Ireland is provided in Annexure 8 for reference.

It was important to include words with regards to top nationalities that the immigrants belong to in Ireland like India, Brazil and Ukraine as many posts and comments would have mentions of these words to refer to immigrants from these nationalities. A total of 1620 posts and 30,562 comments were found to have mentions about immigrants in the reference period.

From the graph below, it is seen that in 2024 there was a big spike in percentage of mentions of immigrants in comments compared to 2023 while, there is a marginal dip in percentage of mentions in posts (while the levels are still very high in the reference period).

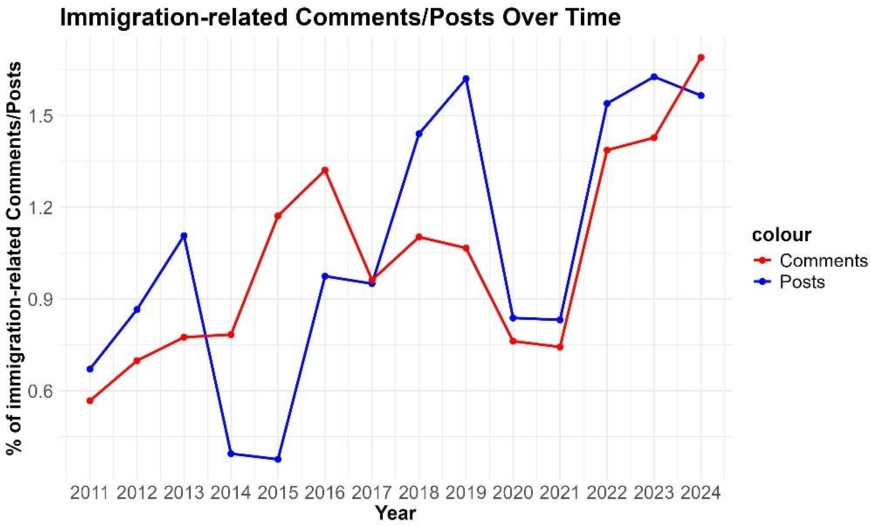

Figure 6: Trend showing immigration-related comments and posts in the reference period

Association with immigration trend in Ireland:

One of the objectives of the study was to understand if there was a relationship between number of immigrants coming to the country and percentage of mentions about immigration in Reddit posts and comments in a year. The total number of new immigrants (refer Table 1) coming to the country provided by the CSO was used for this purpose.

It was observed that the number of new immigrants entering the country had substantially increased since 2022. Most of the new immigrants are from Ukraine that are given refugee status under the temporary protection directive. The table also highlights the number of refugees out of the total number of immigrants which is provided by The UN Refugee Agency. It is noted that 74% of the immigrants were refugees and out of that 80% were Ukrainian refugees.

The scatter plot below demonstrates the association between the number of new immigrants and the percentage of mentions of immigration in Reddit posts and comments during the reference period. The Pearson correlation coefficient indicates a moderate to strong association, with values of 0.71 for posts and 0.77 for comments.

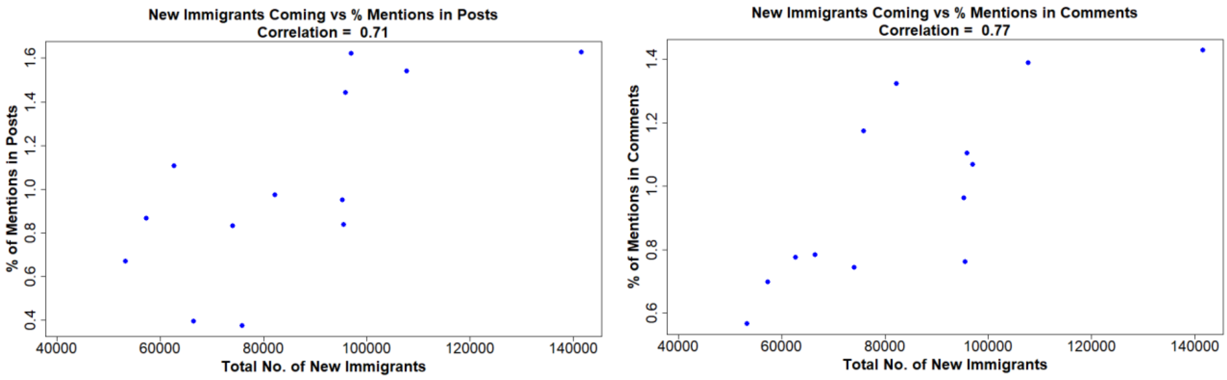

Figure 7: Scatter plot between new immigrants and % of mentions about immigrants in Reddit posts and comments

To further understand this relationship, a linear regression model was applied. The dependent variable was the percentage of mentions about immigrants in posts and comments (analysed separately), and the independent variable was the total number of new immigrants (in multiples of 1,000) entering Ireland during the reference period. The number of new immigrants coming to the country was rescaled to multiples of 1000 for ease of interpretation of results.

Both the bivariate model and a full model with the Year fixed effect was performed. The bivariate model shows a clear association between the two variables. Specifically, for every 1,000 new immigrants, there is a statistically significant increase of 0.0125 percentage points in mentions in posts and 0.001 percentage points in comments. However, when controlling for the year fixed effect, the full model shows no significant relationship between the variables.

Table 2: Regression model estimates with % of mentions in posts and comments as DV and new immigrants as IV

Dependent variable	**% of mentions in posts**		**% of mentions in comments**	
Model number	*Model 1a*	*Model 1b*	*Model 2a*	*Model 2b*
Intercept	-0.051989 (0.332317)	-2.968779 (96.137896)	0.213918 (0.201165)	33.468884 (57.241000)
Tot_immi_r	0.012588** (0.003769)	0.012387 (0.007714)	0.009038** (0.002282)	0.011329* (0.004593)
Year FE		0.001455 (0.047943)		-0.016584 (0.028545)
Adjusted R^2	0.46	0.40	0.5504	0.5216

*Sig. codes: *** p < 0.001, ** p < 0.01, * p < 0.05, + p < 0.1*

Next, for the county level analysis, the data for annual estimated net migration[7] per 1000 of average population for the year-range 2011-2016 and 2016-2022 for each of the 26 counties in Ireland was taken from the CSO website[8]. The CSO doesn't provide data on immigrants at the county level and thus this data was taken as a proxy measure to understand the change in trend of immigration as this also takes into account the number of people emigrating from the county. It is therefore a more reliable measure to understand the increase in immigrants in the country.

To understand the association between annual estimated net migration per 1000 of average population and percentage of mentions about immigrants in both posts and comments, the percentage of mentions about immigrants in both posts and comments was aggregated for each county for the year-range 2010-2016 and 2016-2024. This was then mapped with the net migration data to understand if there was a relationship.

The below scatter plot shows that the net migration is not correlated or associated with percentage of mentions about immigration both in posts and comments indicating no county level variations.

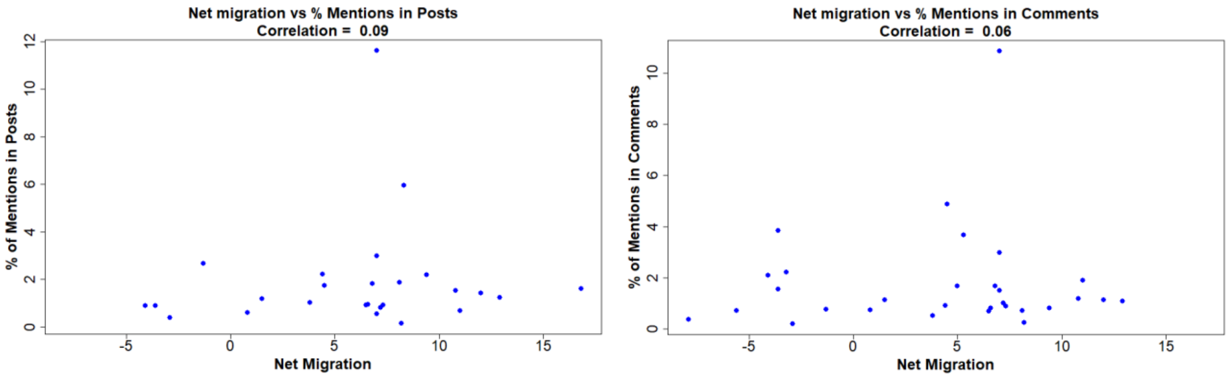

Figure 8: Scatter plot between net migration and % of mentions about immigration in Reddit posts and comments

Further analysis explored the relationship between state-dependent immigrants, such as refugees, and the percentage of mentions about immigrants in posts and comments. The Integration Minister provided county-level data on the number of refugees housed in each county. The analysis focused on data from 2022 onwards, as most refugees, particularly Ukrainians, arrived during this period. The scatter plot below indicates no significant association between the number of refugees housed in each county and the percentage of mentions about immigrants in posts and comments. The Pearson correlation coefficient also supports this finding.

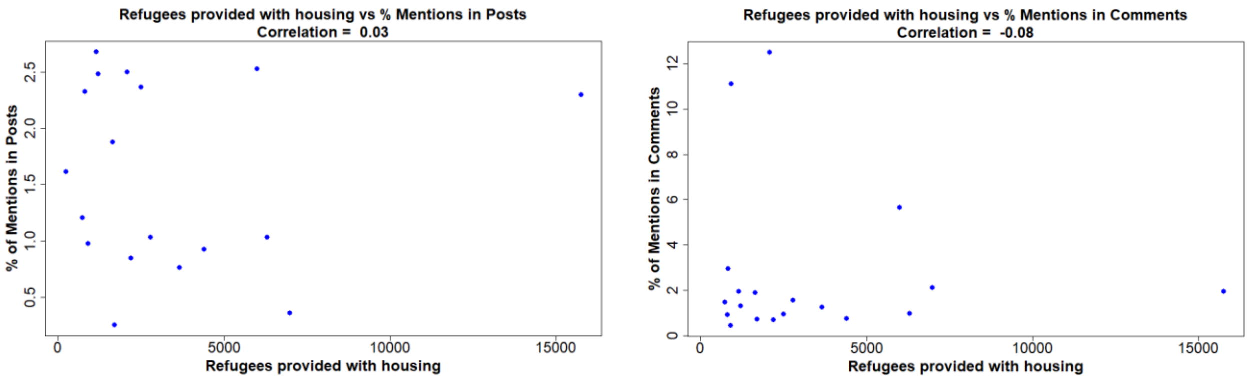

Figure 9: Scatter plot between number of refugees housed in each county and % of mentions in posts and comments

[7] [Background Notes Population and Migration Estimates, April 2023 - Central Statistics Office](#)
[8] https://ws.cso.ie/public/api.restful/PxStat.Data.Cube_API.ReadDataset/F1009/XLSX/2007/en

Lastly, in county-level analysis, relationship between Ukrainians entering the counties and percentage of mentions about immigrants in posts and comments was explored. The data on Ukrainians entering the county was gathered from the Ukraine Hub portal maintained by the CSO. The data for percentage of mentions about immigrants in posts and comments was aggregated for each county and all cases from 2022 onwards were considered as most of the Ukrainians entered the country 2022 onwards.

Similar to the previous analyses, no significant association was found between the number of Ukrainians coming into the country and data from percentage of mentions about immigrants in posts and comments, with a very low Pearson correlation value indicating a weak relationship.

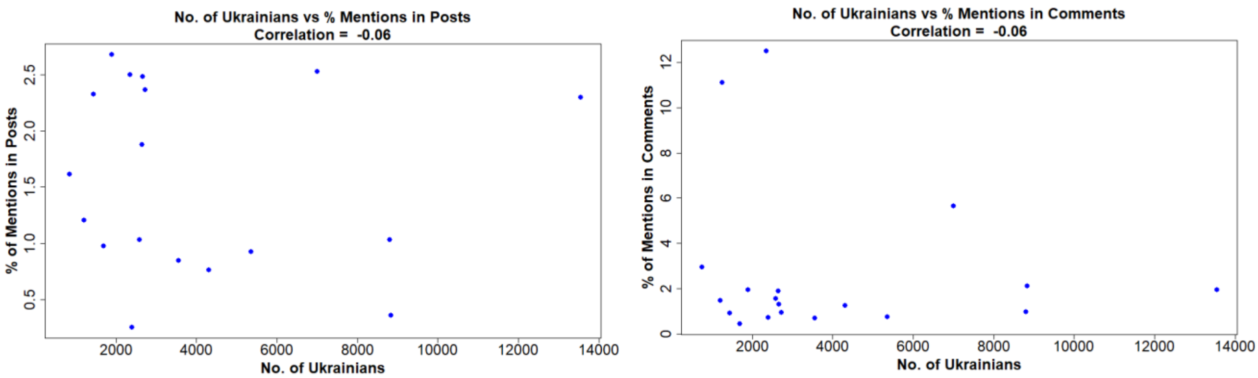

Figure 10: Scatter plot between no. of Ukrainians entering the country and % mentions in posts and comments

3.3 Sentiment analysis:

In this analysis, the sentiment associated with Reddit posts and comments about immigrants in Ireland during the reference period was assessed using the Lexicoder Sentiment Dictionary.

The findings from this analysis indicate that the net sentiment of both posts and comments concerning immigrants has trended increasingly negative over the reference period, with a sharper decline observed in comments after 2020. This trend contrasts with the rising percentage of mentions about immigrants in posts and comments, suggesting an interesting dynamic where increased discourse on immigration is accompanied by a growing negativity in sentiment. Given the significant increase in the number of immigrants, particularly state-dependent immigrants like refugees and asylum seekers, this provides ground to carry out research to explore possible associations between the increase in immigrants and the negative sentiment trends.

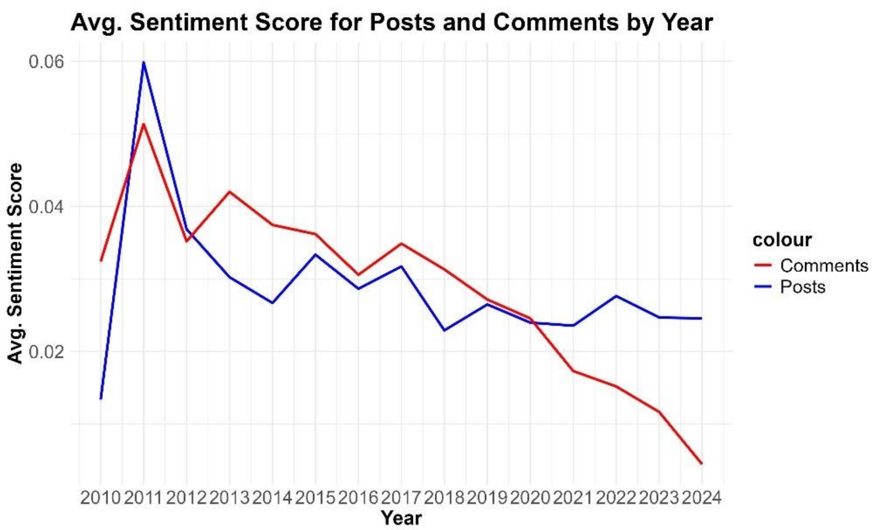

Figure 11: Net sentiment trend for Reddit posts and comments in the reference period

Association with immigration trend in Ireland:

The analysis aimed to explore whether there is a relationship between the number of immigrants entering Ireland and the net sentiment score of Reddit posts and comments during the reference period. The scatter plot below illustrates a moderately negative association between the number of new immigrants and the net sentiment score for Reddit posts (Pearson correlation coefficient = -0.25) and a stronger negative association for Reddit comments (Pearson correlation coefficient = -0.72). These negative correlations suggest that as the number of new immigrants increases, the net sentiment score related to immigrants in posts and comments tends to decline, indicating a shift toward more negative sentiment.

This is an interesting finding as the sentiment score is dependent on the usage of words it implies more negative words in the context of immigration is being used on Reddit.

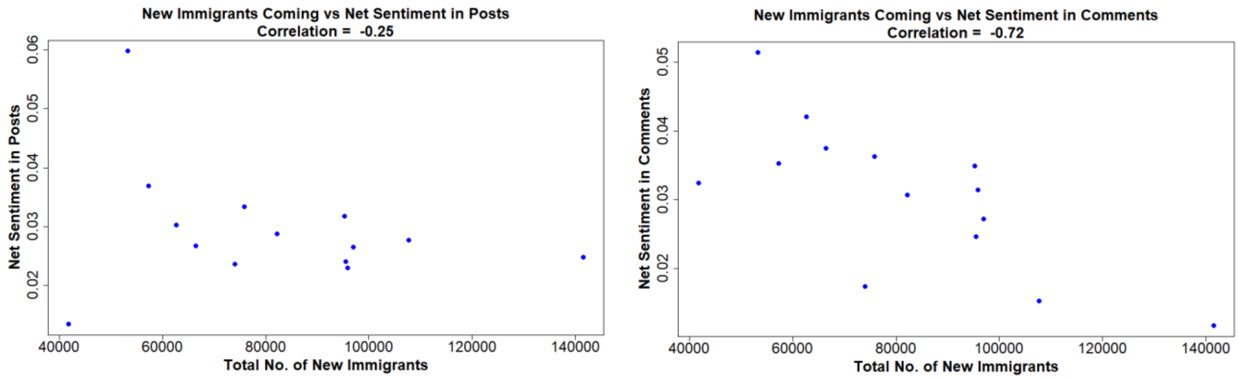

Figure 12: Scatter plot between new immigrants and net sentiment for posts and comments

A linear regression analysis was conducted to further investigate this relationship. The regression model, with the net sentiment score for posts as the dependent variable and the total number of new immigrants entering the country as the independent variable, revealed no significant relationship in both the bivariate model and the full model (which included year as a control variable). However, the regression analysis of comments showed that the number of new immigrants is strongly associated with a decrease in the net sentiment score. Specifically, the bivariate model indicates that for every 1,000 new immigrants entering the country, the net sentiment score in comments decreases (red tending negative) by 0.0003 points.

Table 3: Regression model showing relationship between new immigrants and net sentiment

Dependent variable	**Net sentiment in posts**		**Net sentiment in comments**	
Model number	*Model 1a*	*Model 1b*	*Model 2a*	*Model 2b*
Intercept	3.740e-02** (9.605e-03)	3.1599763 (3.0092074)	5.509e-02*** (7.059e-03)	5.434e+00** (1.655e+00)
Tot_immi_r	-9.917e-05 (1.121e-04)	0.0001233 (0.0002418)	-3.000e-04** (8.238e-05)	8.327e-05 (1.330e-04)
Year FE		-0.0015576 (0.0015010)		-2.683e-03** (8.255e-04)
Adjusted R^2	-0.017	-0.01054	0.4854	0.7136

*Sig. codes: *** p < 0.001, ** p < 0.01, * p < 0.05, + p < 0.1*

At the county level, the analysis explored the relationship between net migration (per 1,000 population) in each county and the net sentiment score for both posts and comments. The scatter plots and Pearson correlation coefficients suggest no significant pattern or relationship at the county level.

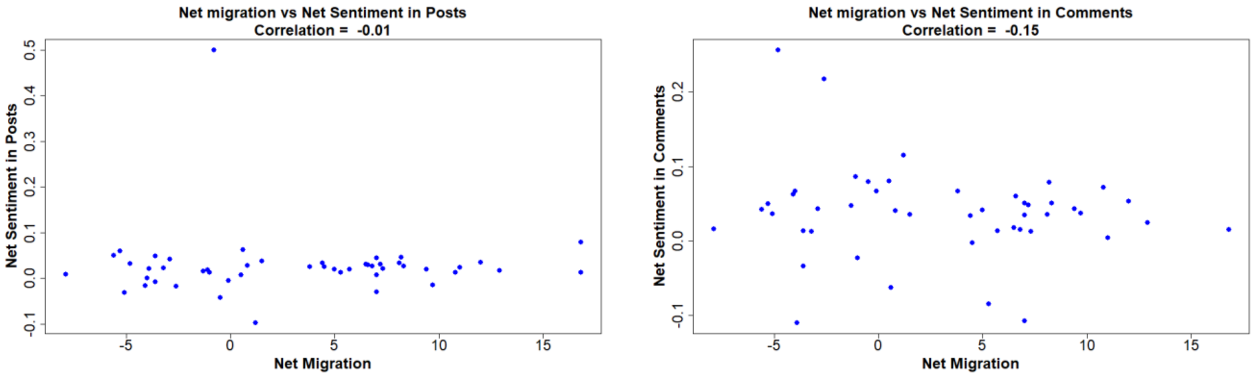

Figure 13: Scatter plot between net migration trend and net sentiment value for posts and comments

Furthermore, no association was found between the number of state-dependent refugees housed in each county and the net sentiment score for both posts and comments. Similarly, there was no significant relationship between the number of Ukrainians entering each county and the net sentiment score for posts and comments, indicating that the sentiment regarding immigrants does not vary significantly across counties.

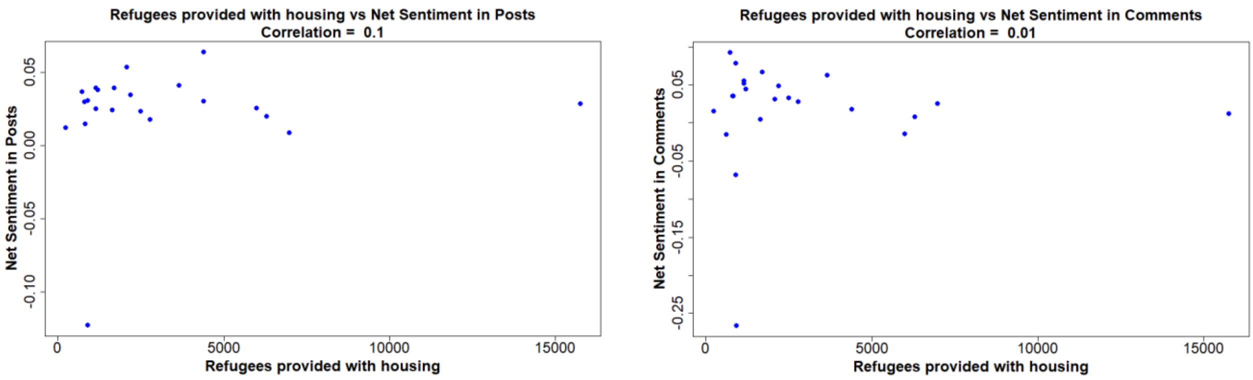

Figure 14: Scatter plot between number of refugees housed and net sentiment

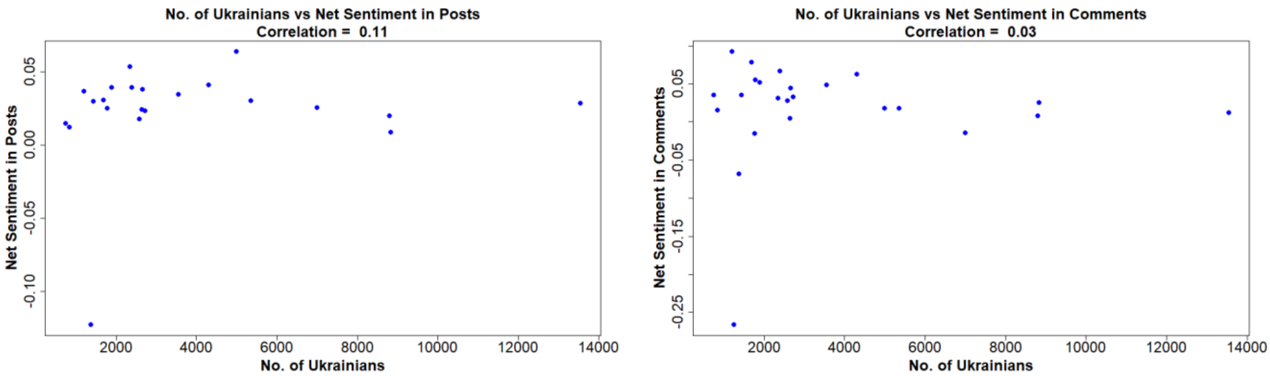

Figure 15: Scatter plot between no. of Ukrainians entering and net sentiment

3.4 Hate speech analysis using supervised learning:

The sentiment analysis showed that the overall sentiment of the contents in both posts and comments related to immigrants were tending towards negative direction. However, it does not necessarily show the level of hate or spite towards the immigrants. It could just be a matter of use of more negative words which the Lexicoder Sentiment Dictionary classifies as negative. To get a better understanding of the hate content towards immigrants, a supervised machine learning model was developed which classified a content as hate speech or not in context of immigrants.

The analysis focused on 1,620 posts related to immigration. A random sample of 500 posts (approximately 30%) was manually labeled by humans and used to train two different machine learning models: Naïve Bayes and Support Vector Machine (SVM). Finally, the model that better performed in classifying the posts as hate speech or not in context of immigrants was selected to predict the labels for all the remaining posts.

The performance of the classifiers is evaluated using a confusion matrix and other performance statistics computed based on the confusion matrix.

3.4.1 Testing Naïve Bayes Model:

The below table shows the confusion matrix and other statistics of the Naïve Bayes model which helps understand the performance of the model for both test and validation data:

Table 4: Confusion matrix and other performance statistics for Naive Bayes model

Naïve Bayes Model	**Test data performance**		**Validation data performance**	
	Reference		***Reference***	
Prediction:	*No*	*Yes*	*No*	*Yes*
No	73	22	17	8
Yes	0	0	0	0
Statistics:				
Accuracy	0.7684		0.68	
95% CI	(0.6706, 0.8488)		(0.465, 0.8505)	
No Information Rate	0.7684		0.68	
P-Value [Acc > NIR]	0.5569		0.59428	
Kappa	0		0	
Mcnemar's Test P-Value	7.562e-06		0.01333	
Sensitivity	1.0000		1.0000	
Specificity	0.0000		0.0000	
Pos Pred Value	0.7684		0.6800	
Neg Pred Value	NaN		NaN	
Precision	0.7684		0.6800	
Recall	1.0000		1.0000	
F1	0.8690		0.8095	
Prevalence	0.7684		0.6800	
Detection Rate	0.7684		0.6800	
Detection Prevalence	1.0000		1.0000	
Balanced Accuracy	0.5000		0.5000	
'Positive' Class	No		No	

The performance of the model on the test data gave a prediction accuracy of 0.76 or 76% which was moderate. The model was then retrained using the “best” parameters settings on all the labelled data (except the validation set) and then the model performance on the validation data was estimated. There was a dip in the prediction accuracy to 68%. The sensitivity and specificity scores indicated that while the model was effective at predicting the "No" class, it struggled to correctly predict the "Yes" class, resulting in a suboptimal model for identifying hate speech. Some more effort on fine tuning could have been done but the model still would not do a good job in predicting the two classes with a decent amount of accuracy.

3.4.2 Testing Support Vector Machine:

Next, SVM was trained on the same training data as the one that was used to train Naïve Bayes model. The below table shows the confusion matrix and other statistics of the model:

Table 5: Confusion matrix and other performance statistics for SVM model

SVM Model	**Test data performance**		**Validation data performance**	
	Reference		***Reference***	
Prediction:	*No*	*Yes*	*No*	*Yes*
No	73	18	16	5
Yes	0	4	1	3
Statistics:				
Accuracy	0.8105		0.76	
95% CI	(0.7172, 0.8837)		(0.5487, 0.9064)	
No Information Rate	0.7684		0.68	
P-Value [Acc > NIR]	0.199		0.2657	
Kappa	0.2546		0.3644	
Mcnemar's Test P-Value	6.151e-05		0.2207	
Sensitivity	1.0000		0.9412	
Specificity	0.1818		0.3750	
Pos Pred Value	0.8022		0.7619	
Neg Pred Value	1.0000		0.7500	
Precision	0.8022		0.7619	
Recall	1.0000		0.9412	
F1	0.8902		0.8421	
Prevalence	0.7684		0.6800	
Detection Rate	0.7684		0.6400	
Detection Prevalence	0.9579		0.8400	
Balanced Accuracy	0.5909		0.6581	
'Positive' Class	No		No	

The SVM classifier achieved an accuracy of 0.815 or 81.5%, with a 95% confidence interval ranging from 54.9% to 90.5% which was much improved than the Naïve Bayes classifier. The high sensitivity (94.1%) suggested that the model is particularly good at identifying the 'No' class, with most 'No' instances correctly predicted. However, the lower specificity (37.5%) indicated difficulty in correctly identifying the 'Yes' class, leading to a higher number of false negatives. The F1-score of 0.8421 which is the harmonic mean of precision and recall for the 'No' class indicated a good balance between precision and recall.

Overall, while the SVM model showed strong performance in predicting the 'No' class, its ability to correctly identify the 'Yes' class could be improved. The model was then used to predict the remaining cases (unlabelled posts). The below graph shows the percentage of hate speech for the Reddit posts in relation to immigrants (predicted by the SVM model) for the reference period. It was seen that the percentage of hate speech in posts significantly spiked after 2022 which was also the time period when a lot of refugees (especially from Ukraine) came into the country. This raises questions about a potential correlation between the increase in refugee arrivals and the rise in hate speech, which will be explored further.

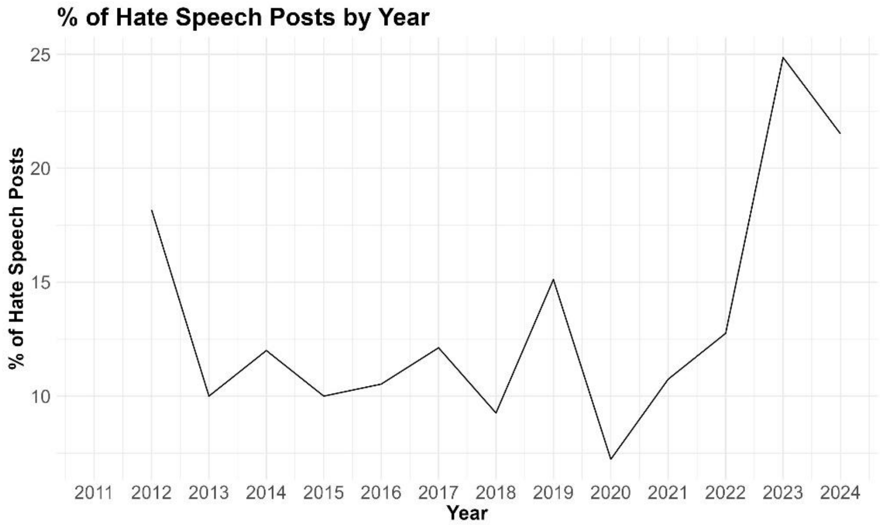

Figure 16: Trend of percentage of hate speech in Reddit posts for the reference period

Association with immigration in Ireland:

To further investigate the relationship between migration and hate speech, scatter plots were generated to examine the correlation between the number of new immigrants entering the country and the percentage of posts labelled as hate speech. The first scatter plot below on the left shows the relation between the number of new immigrants entering the country and percentage of posts labelled with hate speech. It is seen that there is a moderate to low positive association with a Pearson correlation coefficient of 0.47. The second scatter plot below on the right shows the relation between net migration at each county and percentage of posts labelled hate speech. It can be seen that there is a strong positive association with a Pearson correlation coefficient of 0.86.

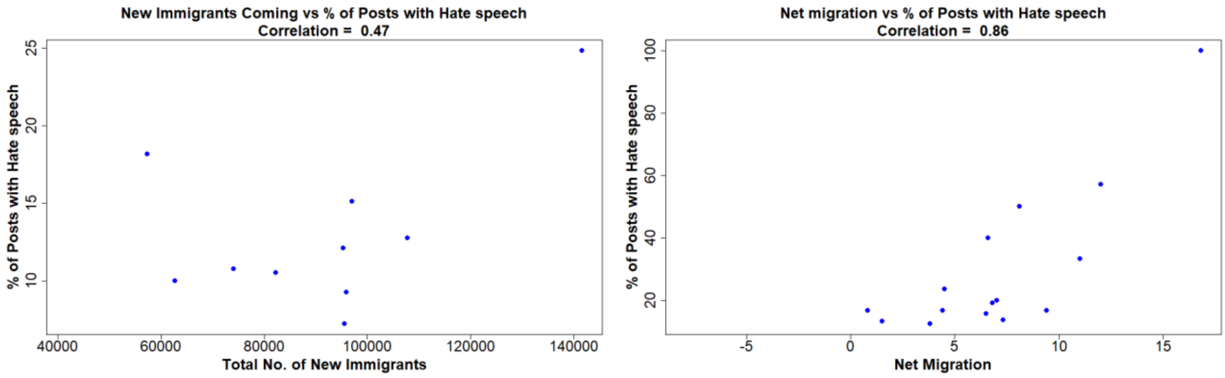

Figure 17: Scatter plot between immigrants and % of posts labelled as hate speech

To further understand the strength of the relationship linear regression models were run with the percentage of posts labelled hate speech as the dependent variable and the number of new immigrants entering the country and net migration at the county level as the independent variables.

It can be observed from the estimates of the model in the table below that there was no association between new immigrants entering the country and percentage of posts labelled hate speech for both bivariate (reduced) and full model (with year fixed effect).

Table 6: Regression model exploring relationship between % of posts labelled as hate speech and new immigrants coming to the country

Dependent variable	% of posts labelled hate speech	
Model number	*Model 1a*	*Model 1b*
Intercept	3.740e-02** (9.605e-03)	3.1599763 (3.0092074)
Tot_immi_r	-9.917e-05 (1.121e-04)	0.0001233 (0.0002418)
Year FE		-0.0015576 (0.0015010)
Adjusted R^2	-0.017	-0.01054

*Sig. codes: *** p < 0.001, ** p < 0.01, * p < 0.05, + p < 0.1*

The second model provided a better indication of the relationship between net migration value for all the counties and percentage of posts labelled hate speech with regards to immigration. It was found that with every one unit increase in net migration for a county for every 1000 population was associated with 5.35% and 6.43% increase in posts being labelled as hate speech with regards to immigration for the reduced bivariate model and full multivariate model (with year-range as additional covariate) respectively. The interaction effect between year-range and net migration value for the counties was also explored to see if there was any effect. However, the coefficients were not significantly different from zero indicating no interaction effect.

Table 7: Regression model exploring relationship between % of posts labelled hate speech and net migration

Dependent variable	% of posts labelled hate speech		
Model number	*Model 2a*	*Model 2b*	*Model 2b*
Intercept	-7.1601 (7.5274)	7.6016 (9.6527)	20.476 (34.085)
Net migration	5.3753*** (0.8433)	6.4334*** (0.9022)	-4.762 (28.355)
Year_range (2016-2022)		-26.1894+ (12.2863)	-39.169 (35.227)
Net migration X Year_range			11.207 (28.371)
Adjusted R^2	0.7254	0.7809	0.7657

*Sig. codes: *** p < 0.001, ** p < 0.01, * p < 0.05, + p < 0.1*

Deep-diving in the county level analysis and exploring the relationship with state-dependent immigrants (especially the refugees), two more hypothesis were tested. The first one being refugees provided with housing in various counties of Ireland and percentage of posts in those counties labelled as hate speech with regards to immigrants. It was observed from the scatter plot and the Pearson correlation coefficient that there is a moderate negative relationship (-0.51). However, the linear regression model showed no statistically significant estimates indicating no association.

Table 8: Regression model with % of posts labelled hate speech as DV and refugees housed as IV

Dependent variable	% of posts labelled hate speech
Model number	*Model 3*
Intercept	53.034626** (11.089043)
Number	-0.003069 (0.001841)
Adjusted R^2	0.165

*Sig. codes: *** p < 0.001, ** p < 0.01, * p < 0.05, + p < 0.1*

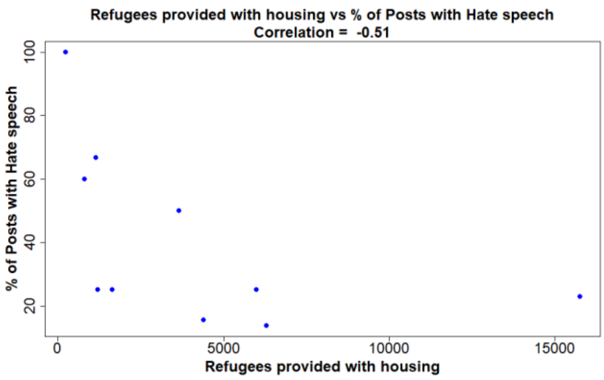

Figure 18: Scatter plot between refugees housed and % of posts labelled hate speech

Next, the second hypothesis where number of Ukrainians entering the counties and percentage of posts in those counties labelled as hate speech with regards to immigration was explored. It was observed from the scatter plot and the Pearson correlation coefficient that there was a moderate negative relationship (-0.62). The linear regression model also indicated a very low association with every addition Ukrainian entering a county, the percentage of posts labelled as hate speech reduced by 0.004% (significant at 10% level). While this is a very low association, it was interesting to note that there was no increase in hate speech incidences with more number of Ukrainians coming into various counties in Ireland.

Table 9: Regression model with % of posts labelled hate speech as DV and no. of Ukrainians entering as IV

Dependent variable	**% of Hate speech in posts**
Model number	*Model 4*
Intercept	61.639231*** (11.991362)
Number	-0.004377+ (0.001951)
Adjusted R^2	0.3095

*Sig. codes: *** p < 0.001, ** p < 0.01, * p < 0.05, + p < 0.1*

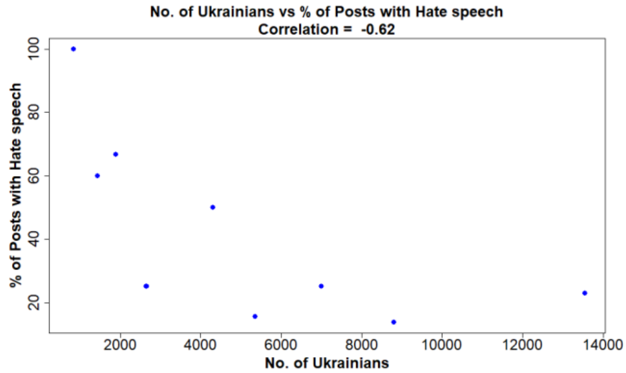

Figure 19: Scatter plot between no. of Ukrainians entering and % of posts labelled hate speech

3.5 Topic modelling:

The sentiment analysis and hate speech prediction gives an indication of people's emotions and sentiments towards immigration and immigrants. To understand what are some of the key areas or topic where people have mostly expressed their views related towards immigrants, topic modelling using Latent Dirichlet Allocation (LDA) method was performed.

As first step all the posts related to immigrants that were labelled as hate speech by the SVM model were filtered out. Then, the texts were broken down into tokens for each post and then necessary text pre-processing (removal of punctuations, special characters, stop words, lowercasing and collocation detection) was done before creating a document-feature matrix. On the clean text dataset, LDA model was then estimated with k=30 topics by fitting the topic model with the Gibbs sampler. The settings were set for a burn in of 100 iterations and a remaining number of 1000 iterations.

As first step in interpreting the LDA results, the topic terms or keywords arranged as per their assignment frequencies was observed followed by trying to understand the semantic coherence of the keywords in each of the topics to see which are related and make sense and which topics are junk. The process of topic interpretation and validation started by examining the topic keywords which are probable tokens under the model – i.e., the tokens most often assigned to a particular topic. The below table shows the top five terms for the estimated topics using the LDA model.

Table 10: Top five terms for all the 30 topics

Topics [1-15]	Topics [16-30]
foreign, foreigners, stay, far, suggest	call, phone, went, anybody, restaurant
foreigner, im, trying, wants, share	house, room, rent, friend, apartment
family, home, without, days, definitely	good, advice, time, near, advance
first, car, card, driving, test	indian, great, good, food, suggestions
give, change, cities, seeing, last	time, town, enough, put, tell
come, soon, came, man, heard	brazilian, english, brazil, portuguese, guys
living, moving, live, move, able	club, events, groups, group, community
money, guy, every, another, stuff	immigrants, immigrant, safe, home, everyone
ukraine, ukrainian, refugees, help, clothes	years, especially, eu, language, social
men, british, road, part, another	due, street, irish, cinema, since
help, thank, anything, kind, hello	india, student, study, year, hey

irish, today, women, small, called	love, feel, since, friends, better
mayo, mayo_day, time, food, irish	work, job, experience, visa, working
immigration, travel, hours, visa, flight	street, bus, happened, left, started
night, college, pub, trip, somewhere	government, irish, housing_crisis, housing, prices

It can be noted that a lot of these collections of terms look coherent. For instance, the 9^{th} topic seems to be about Ukrainian refugees and support provided to them. There are also topics that don't look coherent and could be classified as "junk". Based on the top keywords, some topics look very clear, while others are a bit more ambiguous in their underlying meaning. To get a better sense of what a particular topic is about, a handful of posts were read for some of the topics that were coherent. This qualitative validation process helped in making a better sense of what the topics should be labelled and understanding how well the topic label fits the underlying text.

Besides qualitative validity, predictive validity of the model was also performed on some of the topics. For instance, topic 23 relates to immigration and immigrants. The below graph shows the increasing average yearly topic probability trend of the "immigrant safety" topic for the reference period. This is in line with the increase in percentage of mentions about immigrants over time as a greater number of immigrants enter into the country especially after 2022. This is further evidence that the interpretation of the topic is made correctly.

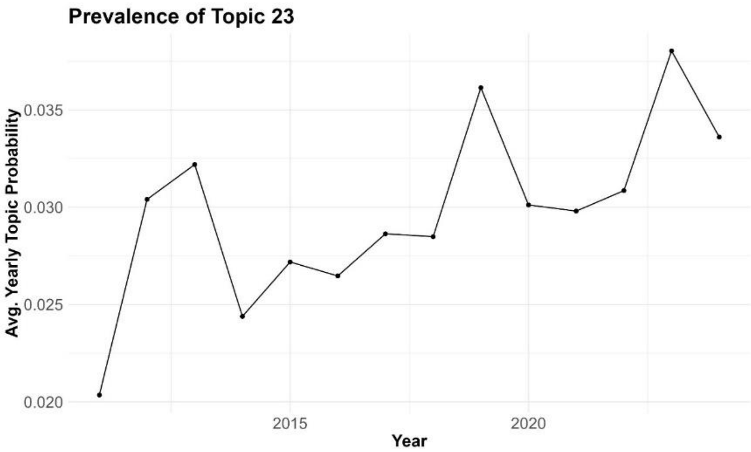

Figure 20: Prevalence of topic 23 over time

Some of the topics that were found from the topic model are about getting a first car (topic 4), able to move and live abroad (topic 7), Ukrainians refugees and support services provided (topic 9), travel visa and immigration enquiry (topic 14), seeking housing (topic 17), Indian food suggestions (topic 19), clubs and event groups (topic 22), immigrant safety (topic 23), Indian student enquiry (topic 26), work and visa requirement (topic 28), housing crisis and government role (topic 30).

It can be seen that topics in relation to immigrants are quite valid and expected like enquiring about visa requirements for work and travel and also concerns about finding a place to stay given the housing crisis. Also, discussions around support provided to Ukrainian refugees and about immigrant safety is quite relevant. It was also interesting to note enquiry topics about Indian food suggestions and study related enquiry by Indian prospective students.

4 Discussion and implications for future work:

The analysis of Reddit content on immigration in Ireland revealed significant trends and relationships between public sentiment online and increasing immigration levels in the country. This section will look at the implications of the study findings and identify potential for future work.

The study found a clear trend of increasing negativity in sentiment towards immigrants, particularly in the post-2020 period. This is evident from both the sentiment analysis, which showed a decline in net sentiment scores over time, and the hate speech analysis, which revealed a notable spike in hate speech incidents after 2022. These findings suggest that the discourse around immigration on social media (especially Reddit) is becoming more negative. It reflects broader concerns and anxieties about immigration. This trend also aligns with the increasing number of mentions about immigrants on Reddit, indicating that as immigration becomes a more crucial topic of discussion, it is increasingly associated with negative sentiments.

The strong negative correlation between the number of new immigrants entering Ireland and the net sentiment value of Reddit comments suggests that as the number of immigrants increases, public sentiment as expressed in comments becomes more negative. The hate speech analysis adds another layer of complexity to the understanding of public sentiment towards immigrants. The increase in percentage of posts labelled as hate speech after 2022, suggests that significant influxes of immigrants can lead to more negative and hateful content. This highlights the importance of monitoring social media content by the platform moderators and identify patterns of ambient racism that Sharm (2018) talks about before racial expressions on the internet are normalised.

It is hoped that the study lays the foundation for future work in detection of online hate speech. The machine learning algorithm can further be improved and fine-tuned to improve prediction accuracy. More work can also be done to develop multi-class supervised classifier which detects different degrees of hate speech (neutral, moderate and extreme). The classifier thus developed could be integrated into social media platforms to automatically monitor and flag content that is likely to be classified as extreme hate speech. By incorporating such a tool, platforms could reduce the spread of ambient forms of racism and other harmful content. The classifier could be fine-tuned and regularly updated to improve its accuracy and effectiveness, ensuring that it can adapt to new forms of language and context used to convey hate speech.

Furthermore, though in recent years, the rise in immigrants is primarily because of arrival of Ukrainian refugees, there was absence of any significant increase in hate speech. It challenges the notion that all large influxes of immigrants would necessarily lead to an increase in negative and hateful content on social media. This may be due to a variety of factors, including public perceptions of Ukrainians as victims of a high-profile conflict, positive media coverage with regards to Ukrainian refugees, or differences in the cultural or social integration of this group compared to other immigrant populations. This finding suggests that public sentiment towards immigrants is not solely driven by the number of immigrants but can also be influenced by the specific context and characteristics of the immigrant population.

Lastly, the Irish government has a key role to play in properly integrating the immigrants (especially the refugees) with the local population to reduce friction. International Organization for Migration (IOM) developed a report recommending government to lead the way in managing the immigrants (IOM, 2006). This report suggests that migration integration policies should be mainstreamed into services like health and education for the economic development of the country which is still applicable in today's context. Further awareness campaigns and movements at the county level should be arranged for wider public sensitisation on the issue of immigration. The government should consider launching public awareness campaigns that highlight the contributions of immigrants, particularly refugees, to Irish society as part of the integration process. By promoting positive narratives, it is possible to counterbalance the negative sentiments and reduce the potential for social friction.

4.1 Study limitations:

The study has several limitations that could have impacted the findings. First, the analysis is based solely on Reddit data, which may not fully reflect the broader public's views. Reddit users are typically younger and more tech-savvy than the general population, which could skew the results. Having a better understanding of the demographic of the Reddit users will help to generalise the findings to the wider Irish population.

Second, the dictionary used for sentiment analysis, while useful, have limitations in understanding the full context of online conversations. The dictionary might misinterpret certain words or phrases out of context. In the issue of context, while efforts were made to improve the machine learning model, detecting subtle and context-dependent hate speech remains difficult. More advanced models, like neural networks or transformers, might be better at understanding these nuances. It is hoped that this study lays the ground for further refinement and testing of models in detection of online hate speech.

Regarding the annotation process, six human annotators were involved, and even though guidelines were provided, some errors in labeling could still occur. To reduce bias, each piece of data was labeled three times, with the majority label chosen as final. However, given that it’s a human work, it is prone to errors.

Lastly, the study focuses only on Reddit posts, starting from 2010. This limits the applicability of the findings to other platforms. The SVM classifier was specifically trained on Reddit data, which might limit its effectiveness when applied elsewhere. Also, only Reddit posts were used to train the model, leaving out comments, which might offer different perspectives. Future studies could also explore models that work better with comment data.

5 Conclusion:

In conclusion, this study highlights the growing negativity in public sentiment towards immigrants in Ireland, particularly in the context of increasing immigration levels. The relationship between immigration and hate speech, while complex, appears to be influenced by both the number of immigrants and contextual factors, with certain types of immigration more closely associated with negative discourse. These findings bring to light the role of social media platforms to put in place strict moderations and robust mechanisms to detect and delete hateful content. Moreover, the role of Irish government in addressing public concerns and promoting positive narratives around immigration, particularly in counties experiencing higher levels of immigration is also important. This will help curb the rise of hate speech on online media and also ensure better social cohesion between immigrants and locals.

6 Bibliography

- Official Journal of the European Union. (2022, 3 21). *INFORMATION FROM EUROPEAN UNION INSTITUTIONS, BODIES, OFFICES .* Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022XC0321(03)

- *Addressing Ireland's Housing Crisis: Urgent Policy Reforms Needed.* (2024, 5). Retrieved from Social Justice Ireland: https://www.socialjustice.ie/article/addressing-irelands-housing-crisis-urgent-policy-reforms-needed

- Benoit, K. K. (2018). quanteda: An R package for the quantitative analysis of textual data. *3*(30), 774. doi:https://doi.org/10.21105/joss.00774

- Benoit, K. K. (2018). Quanteda: An R package for the quantitative analysis of textual data. *3*(30), 774. doi:https://doi.org/10.21105/joss.00774

- Benoit, K., Munger, K., & Spirling, A. (2019, April). Measuring and Explaining Political Sophistication Through Textual Complexity. *American Journal of Political Science, 63*, 491–508. doi:10.1111/ajps.12423

- Central Statistics Office. (2023). *Census of Population 2022 – Summary Results: Migration and Diversity'.* CSO statistical publication. Retrieved from www.cso.ie

- Citizens Information. (2024, July). *International protection terms explained*. Retrieved from https://www.citizensinformation.ie/en/moving-country/asylum-seekers-and-refugees/the-asylum-process-in-ireland/international-protection-terms/

- Deegan, G. (2023). Some asylum seekers decide to leave Clare after blockade at former hotel. Irish Times.

- Disch, W., & Slaymaker, R. (2023). Housing affordability: Ireland in a cross-country context. *ESRI*(164). doi:https://doi.org/10.26504/rs164

- Ekman, M. (2018). Anti-refugee mobilization in social media: The case of soldiers of Odin. *Social Media + Society, 4*(1), 1-11.

- Ekman, M. (2019). Anti-immigration and racist discourse in social media. *European Journal of Communication, 34(6) 606 –618*. doi:10.1177/0267323119886151

- Farkas, J., Schou, J., & Neumayer, C. (2018). Platformed antagonism: Racist discourses on fake Muslim Facebook pages. *Critical Discourse Studies, 15*(5), 463–480.

- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221–233. doi:https://doi.org/10.1037/h0057532

- Grimmer, J. &. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *21*, 267-297. Retrieved from https://doi.org/10.1093/pan/mps028

- Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. *Elsevier*.

- Hopkins, D., & King, G. (2010, January). A Method of Automated Nonparametric Content Analysis for Social Science. *American Journal of Political, 54*(1). doi:10.1111/j.1540-5907.2009.00428.x

- IOM. (2006). *Managing Migration in Ireland: A Social and Economic Analysis.* Retrieved from https://publications.iom.int/books/managing-migration-ireland-social-and-economic-analysis

- Irish Mirror. (2023). *County-by-county breakdown of Ukrainian refugees and asylum seekers in Ireland*. Retrieved from https://www.irishmirror.ie/news/irish-news/politics/refugees-ireland-ukraine-dublin-breakdown-29127043

- Jurafsky, D., & Martin, J. (2008, February). Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. *ResearchGate*.
- Keena, C. (2022). East Wall residents protest against new refugee accommodation. Irish Times.
- Kumar, C., & Donoghue, D. (2023). *Public narratives and attitudes towards refugees and other migrants: Ireland country profile.* London: ODI.
- Lally, C., & Fallon, J. (2023). *Galway hotel fire: Gardaí believe blaze at premises due to house 70 asylum seekers started*. Irish Times.
- Lally, C., & J, H. (2024). 'Fire at Dublin building intended as . Irish Times.
- Larsson, N. (2018). Ljusdalskvinna döms för hets mot folkgrupp – skrev nedsättande om somalier på Facebook [Woman from Ljusdal convicted of hate speech – wrote derogatory comments targeting Somalians on Facebook.]. *HelaHälsingland*. Retrieved from https://www.helahalsingland.se/artikel/ljusdalskvinna-doms-for-hets-mot-folkgrupp-skrev-nedsattande-om-somalier-pa-facebook
- M. Blei, D., Ng, A., & Jordan, M. (2001). Latent Dirichlet Allocation. *Journal of Machine Learning Research, 3*, 601-608.
- Matamoros-Fernández, A. (2017). Platformed racism: The mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube. *Information, Communication and Society, 20*(6), 930–946.
- McGinnity, F. S. (2023). Monitoring report on integration 2022. *ESRI and Department of Children, Equality, Disability, Integration and Youth*. doi:https://doi.org/10.26504/jr3
- Miles, R., & Brown, M. (2003). Racism, 2nd ed. *New York: Routledge*.
- Murphy, K., & Sheridan, A. (2023). Annual report on migration and asylum 2022. (124). doi:https://doi.org/10.26504/sustat124
- National Youth Council of Ireland. (2024, May). *Asylum seekers, refugees, migrants – understanding the Irish immigration system*. Retrieved from https://www.youth.ie/articles/asylum-seekers-refugees-migrants-understanding-immigration-systems/#:~:text=An%20asylum%20seeker%20is%20a,no%20longer%20an%20asylum%20seeker.
- Nikunen, K. (2018). From irony to solidarity: Affective practice and social media activism. *Studies of Transition States and Societies, 10*(2), 10-21.
- Peter McVerry Trust. (2024, 5). Retrieved from https://homelessnessinireland.ie/
- Porter, M.F. (1980). *New models in probabilistic information retrieval.* British Library. London: British Library Research and Development Report. Retrieved from https://tartarus.org/martin/PorterStemmer/def.txt
- Sarker, I., Kayes, A., Badsha, S., Alqahtani, H., Watters, P., & A, N. (2020). Cybersecurity data science: an overview from machine learning perspective. *J Big Data, 7*(1), 1–29. doi:10.1186/s40537-020-00318-5
- Schütze, Hinrich, M., & Christopher, D. (1999). Foundations of Statistical Natural Language Proccesing. *The MIT Press Cambridge*.
- Sharma, S. (2018). Affect and the attention economy of online racism. *Proceedings of the 67th annual conference of the international communication association.* Prague.
- Slater, S., & Lucey, A. (2023). Protests at Rosslare Europort and in Killarney over international protection centres. Irish Times.

- The Housing Commission. (2024). *Report of The Housing Commission.* Retrieved from https://www.gov.ie/pdf/?file=https://assets.gov.ie/294018/e1aae1ed-07c4-473d-811e-3426756321ee.pdf#page=null
- Törnberg, A., & Wahlström, M. (2018). Unveiling the radical right online: Exploring framing and identity in an online anti-immigrant discussion group. *Sociologisk Forskning, 55*(2-3). doi:https://doi.org/10.37062/sf.55.18193
- Wahlström, M., & Törnberg, N. (2019, April 4). Social media mechanisms for right-wing political violence in the 21st century: Discursive opportunities, group dynamics, and co-ordination. *Terrorism and Political Violence*. doi:10.1080/09546553.2019.1586676
- Wallach, H. M., Mimno, D., & McCallum, A. (2009). Rethinking LDA: why priors matter. 1973 - 1981.
- Young, L., & Stuart, S. (2012). Affective News: The Automated Coding of Sentiment in Political Texts. *ResearchGate, 29*(2), 205-231. doi:10.1080/10584609.2012.671234

List of annexures

Annexure 1: Subreddit handles of all the 26 counties

Table 11: List of counties with their subreddit handles

County	Subreddit	County	Subreddit
Carlow	r/Carlow	**Longford**	r/Longford
Cavan	r/Cavan	**Louth**	r/countylouth
Clare	r/Clare	**Mayo**	r/mayo
Cork	r/Cork	**Meath**	r/meath
Donegal	r/Donegal	**Monaghan**	r/Monaghan
Dublin	r/dublin	**Offaly**	r/offaly
Galway	r/galway	**Roscommon**	r/Roscommon
Kerry	r/Kerry	**Sligo**	r/Sligo
Kildare	r/Kildare	**Tipperary**	r/Tipperary
Kilkenny	r/Kilkenny	**Waterford**	r/Waterford
Laois	r/Laois	**Westmeath**	r/westmeath
Leitrim	r/Leitrim	**Wexford**	r/CountyWexford
Limerick	r/Limerick	**Wicklow**	r/Wicklow

Annexure 2: Net migration figures (for every 1000 population) for each county

Table 12: Net migration figures for each county

Period	County	Net_mig	Period	County	Net_mig
2011 - 2016	Carlow	-1	2011 - 2016	Longford	1.2
2016 - 2022	Carlow	8.1	2016 - 2022	Longford	16.8
2011 - 2016	Cavan	-0.1	2011 - 2016	Louth	0.6
2016 - 2022	Cavan	5.7	2016 - 2022	Louth	7.4
2011 - 2016	Clare	-3.9	2011 - 2016	Mayo	-4.8
2016 - 2022	Clare	8.3	2016 - 2022	Mayo	6.6
2011 - 2016	Cork	0.8	2011 - 2016	Meath	-0.5
2016 - 2022	Cork	7.3	2016 - 2022	Meath	12.9
2011 - 2016	Donegal	-7.9	2011 - 2016	Monaghan	-4
2016 - 2022	Donegal	4.5	2016 - 2022	Monaghan	5
2011 - 2016	Dublin	1.5	2011 - 2016	Offaly	-5.1
2016 - 2022	Dublin	6.8	2016 - 2022	Offaly	5.3
2011 - 2016	Galway	-2.9	2011 - 2016	Roscommon	-3.6
2016 - 2022	Galway	6.5	2016 - 2022	Roscommon	10.8
2011 - 2016	Kerry	-1.3	2011 - 2016	Sligo	-5.6
2016 - 2022	Kerry	7	2016 - 2022	Sligo	8.2
2011 - 2016	Kildare	-0.8	2011 - 2016	Tipperary	-5.3
2016 - 2022	Kildare	9.7	2016 - 2022	Tipperary	4.4
2011 - 2016	Kilkenny	0.5	2011 - 2016	Waterford	-3.2
2016 - 2022	Kilkenny	3.8	2016 - 2022	Waterford	11
2011 - 2016	Laois	0.6	2011 - 2016	Westmeath	-2.6
2016 - 2022	Laois	7	2016 - 2022	Westmeath	7
2011 - 2016	Leitrim	-3.6	2011 - 2016	Wexford	-1.6
2016 - 2022	Leitrim	12	2016 - 2022	Wexford	10.8
2011 - 2016	Limerick	-4.1	2011 - 2016	Wicklow	-1.1
2016 - 2022	Limerick	7.2	2016 - 2022	Wicklow	9.4

Annexure 3: Total number of refugees provided with accommodation

Table 13: Total number of refugees provided with accommodation in each county

County	Total refugees housed	County	Total refugees housed
Carlow	807	Longford	242
Cavan	1162	Louth	1350
Clare	4394	Mayo	3660
Cork	6304	Meath	2785
Donegal	5991	Monaghan	831
Dublin	15750	Offaly	919
Galway	4409	Roscommon	738
Kerry	6986	Sligo	1698
Kildare	615	Tipperary	1218
Kilkenny	901	Waterford	1643
Laois	911	Westmeath	2089
Leitrim	1153	Wexford	1345
Limerick	2203	Wicklow	2505

Annexure 4: Total cumulative number of Ukrainian refugees entering each county

Table 14: Total cumulative number of Ukrainian refugees entering each county as of 9th Oct 2023

County	Number	County	Number
Clare	5006	Leitrim	1898
Cork	8808	Laois	1385
Cavan	1787	Meath	2593
Carlow	1448	Monaghan	755
Donegal	7006	Mayo	4316
Dublin	13543	Offaly	1251
Galway	5355	Roscommon	1212
Kildare	1772	Sligo	2400
Kilkenny	1702	Tipperary	2658
Kerry	8839	Waterford	2650
Longford	853	Wicklow	2723
Louth	2236	Westmeath	2347
Limerick	3562	Wexford	3471

Annexure 5: Detailed steps followed for text analysis

Acquiring the corpus:

The unit of analysis is a Reddit post or comment thus, the corpus consisted of all comments or posts scraped from the period 2010 to May 2024 from various subreddits relevant to the Irish context.

Pre-processing the text:

The preprocessing of textual corpus involved several key steps to ensure the data was ready for analysis. First, tokenisation was performed to split documents into individual tokens. While the preprocessing steps leads to a substantial reduction in text, scholars have argued that this simple representation of texts retains all the valuable information and interesting properties to extract meaning from it (Hopkins & King, 2010). Next, stemming was applied using the Porter stemming algorithm (Porter, M.F., 1980) to reduce words to their root forms, thereby minimising the number of unique words and the overall dimensionality of the data. The cleaning process then involved removing special symbols, signs, punctuation marks, numbers, and extra spaces, as well as converting all tokens to lowercase to maintain uniformity.

Stop word removal was then conducted to discard very common words that do not contribute significant meaning, using Quanteda's English language stop-words function along with additional context-specific stop words such as "Ireland," "county," "country," "Dublin," "Irish," "people," and "Guinness." Finally, collocation conversion was applied to identify and convert common sequences of tokens into multi-word expressions (Schütze, Hinrich, & Christopher, 1999).

Creating the DFM:

After pre-processing of the text, a document-feature matrix (DFM) is created. This matrix is a collection of vectors representing the number of times a word occurs in a document. The structure of the DFM is such that the rows represent documents in this case unique ID of the Reddit posts and comments and columns represent the features or the words and each cell have the frequency of a given feature in a given document. The DFM serves as the input for further analysis.

Analysis of textual data:

Various quantitative and statistical procedures are applied to the DFM for text analysis. Textual statistics like frequency analysis which identifies the most commonly occurring features and summarises them as frequency distribution is done. It helps highlight the top features of words present in the corpus. Also, word clouds are created which visually help get a sense of the text.

Annexure 6: Developing the training set and supervised learning process flow

Preparing the training set for machine learning models involved careful labelling to ensure its quality and relevance. This process was carried out through crowdsourcing and expert annotation. Following Krippendorff's (2013) guidelines, a comprehensive coding guideline was developed to assist human annotators in categorising Reddit posts as either hate speech or not. This guideline included examples to illustrate the categorisation process. To ensure consistency, a workshop was conducted with all six human annotators to orient them on the coding guidelines, ensuring everyone was on the same page. Multiple coders annotated each document to maintain unbiased and consistent labelling. Discrepancies in labelling were resolved using a majority rule, where the most common label was selected. Two groups of three human annotators each coded 250 Reddit posts, resulting in a total of 500 posts being labelled. Documents with completely divergent labels were excluded to maintain dataset integrity. Adhering to these practices ensured that the training set was robust and well-prepared for effective model training and evaluation.

The coding guideline that was shared with the human labellers is provided below:

Guideline Document for Human Labellers:
Categorising Social Media Comments on Immigration

Introduction:

Welcome to the team and thank you for helping me create a training set by labelling comments from social media (Reddit) into specific categories of hate speech. Given the fact there are multiple people working on the assignment, it is imperative to bring some standardisation measure in how we want the labelling to be done. Hopefully, this document will provide the clear instructions and necessary examples to ensure consistency and accuracy in the labelling process.

Categories of hate speech:

Classification of each comment into one of the following categories:

- Hate speech
- Not hate speech

1. Not hate speech:
Comments that include general discussions and mentions about immigrants without any negative connotation. These comments are factual, informative, or neutral in tone.

Examples:

- "Immigrants have significantly contributed to our economy."
- "I love Indian food. You should definitely try Hyderabadi Kitchen in Dublin."
- "Immigration policies have been a topic of debate for couple of years now in Ireland."

2. **Hate speech:**
Comments that include stereotypes or minor accusations towards immigrants. These comments may be insensitive or perpetuate generalisations but do not incite violence or contain direct threats.

Examples:

- "Immigrants are taking all the jobs."
- "They should adopt the local culture if they want to live here."
- "Most immigrants just come here to exploit our welfare system."

Comments that include direct attacks, threats, or highly inflammatory language towards immigrants. These comments are aggressive, hostile, and may incite violence or hatred.

Examples:

- "Immigrants should be deported immediately!"
- "They are a threat to our society and should be eliminated."
- "I can't stand those people; they should go back to where they came from."

Steps for labelling:

1. **Read the comment:** Carefully read the entire comment present in the "body" column in the Excel file to understand its context and content.
2. **Identify the category:** Determine which category (hate speech or not) best fits the comment based on the descriptions and examples provided.
3. **Select the category:** Assign the appropriate label to the comment in the "hate speech category" column.
4. **Move to the next comment:** Proceed to the next comment and repeat the process.

Additional tip for accurate labelling:

Ensure you fully understand the context of the comment before labelling it, as the hate speech category is defined by the sentiment of the commenter towards immigrants. A comment might have a lot of cuss words and negative tonality but still express positive sentiment towards immigrants, making it not a hate speech.

For instance: "Some far-right morons have created an anti-immigrant atmosphere in the country. We, as Irish people, absolutely condemn their behaviour and bigotry. What a bunch of fucking losers bringing disgrace to the nation."

Conclusion:

I hope this document assists you in the labelling process. Your work is crucial in building a reliable and accurate training set for detecting hate speech on social media. This training set will be used in a supervised learning model to make predictions for unseen and unlabelled data/comments.

I sincerely thank you for your attention to detail and dedication to this task. Your efforts will contribute significantly to my research. Please don't hesitate to reach out if you need any clarifications.

~~ x ~~

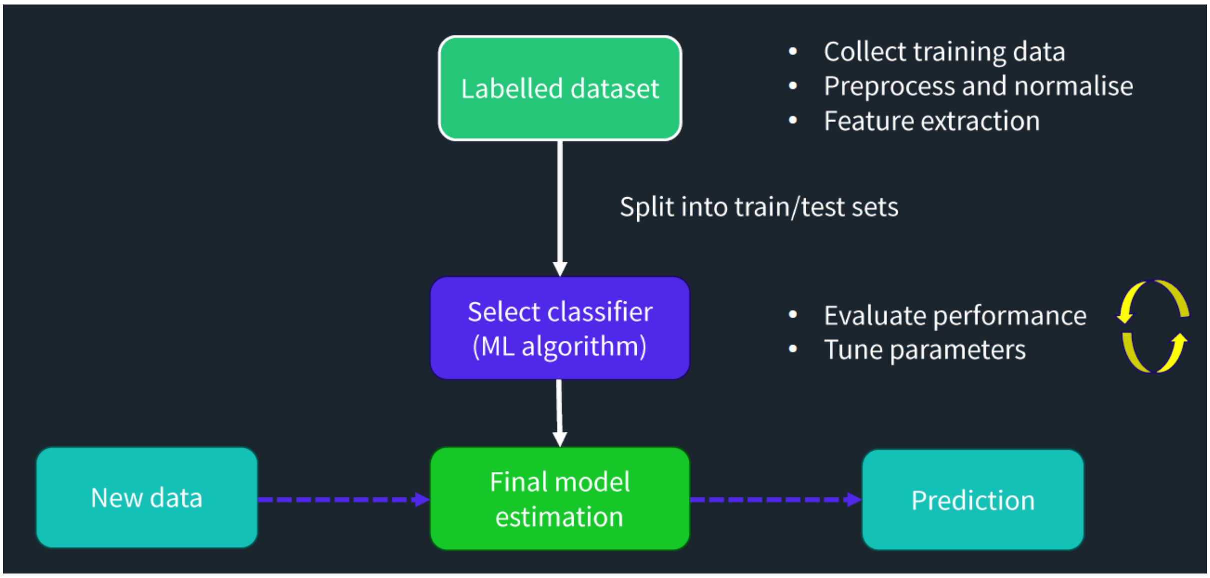

Figure 21: Supervised machine learning process flowchart

Annexure 7: Steps to using LDA for topic modelling:

- **Text preprocessing**: Using a minimalist approach, the text preprocessing was done which involved basic cleaning and stop words removal.
- **Hyperparameter optimisation**: Adopting Wallach's (2009) suggestion, hyperparameter optimisation is conducted to enhance model performance. This included tuning parameters such as the Dirichlet priors for document-topic and topic-word distributions.
- **Determining the number of topics (k)**: Following best practices in qualitative analysis and going by an intuitive judgement, 30 topics were selected.
- **Model validation**: Model validity was ensured through coherence and interpretability metrics, thereby evaluating the quality and relevance of the identified topics. Predictive validity of the model was done by taking a specific topic and looking at the frequency of its occurrence over time. Further qualitative assessment was also done to validate the relevance and accuracy of the topics in relation to the research objectives.

Results interpretation: Finally, the topics were interpreted within the context of the study, linking them to the overarching research objectives.

Annexure 8: List of words used to filter immigrant cases in Reddit posts and comments

Table 15: List of words used to filter immigrant cases in Reddit posts and comments

Term Group	Terms used to filter
Immigration	immigration, immigrations, immigrant, immigrants
Migration	migration, migrations, migrant, migrants
Foreign	foreign, foreigners
Non-Irish/Non-EU	non-Irish, non-EU, non-Europeans
Country-Specific	India, Indian, Indians, Brazil, Brazilian, Brazilians, Ukraine, Ukrainian, Ukrainians
Asylum/Refugees	asylum, refugee, refugees