

Received 7 July 2023, accepted 28 July 2023, date of publication 7 August 2023, date of current version 24 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3303105

## RESEARCH ARTICLE

# InÉire: An Interpretable NLP Pipeline Summarizing Inclusive Policy Making Concerning Migrants in Ireland

AREFEH KAZEMI<sup>1</sup>, ARJUMAND YOUNUS<sup>2</sup>, MINGYEONG JEON<sup>3</sup>,  
M. ATIF QURESHI<sup>3</sup>, AND SIMON CATON<sup>1</sup>

<sup>1</sup>School of Computer Science, University College Dublin, Dublin 4, D04 V1W8 Ireland

<sup>2</sup>School of Sociology, University College Dublin, Dublin 4, D04 V1W8 Ireland

<sup>3</sup>ADAPT Centre, eXplainable Analytics Group, Faculty of Business, Technological University Dublin, Dublin 2, D02 HW71 Ireland

Corresponding author: Arefeh Kazemi (arefeh.kazemi@ucd.ie)

This work was supported in part by the Irish Research Council under Award COALESCE/2021/112, and in part by the Science Foundation Ireland with the ADAPT SFI Research Centre under Grant 13/RC/2106\_P2.

**ABSTRACT** Reaching marginal and other migrant communities to elicit their political views and opinions is a well-known challenge. Social media has enabled a certain amount of online activism and participation, especially in societies with abundant multicultural identities. However, it can be quite challenging to isolate the voice of the migrant in English-speaking countries, especially with an abundance of content in English on social media. In this paper, we pursue a case study of Ireland's Twitter landscape, specifically migrant and native activists. We present a methodology that can accurately (> 80%) isolate the Irish migrant voice with as little as 25 English tweets without relying on user metadata and using simple, highly explainable, out-of-the-box machine learning methods. Using this, we distil (via sentiment analysis) polarities of views, segment (via BERT-based topic modelling) and summarise (via ChatGPT) differentiated views in a consumable manner for policymakers. Our approach enables policymakers to further their understanding of multicultural communities and use this to inform their decision-making processes.

**INDEX TERMS** Natural language processing, Ireland, migrant, Twitter, summarization, policy making.

## I. INTRODUCTION

Social media has become an essential day-to-day communication platform [1], [2] that can be employed to study specific social contexts and processes through the lens of the user-generated content it encompasses. While having a multi-faceted role as a communications platform, it has also become a platform for knowledge sharing, activism, and journalism. The scale of the underpinning network structure allows social media researchers to answer questions that perhaps more traditional methods cannot, for example, when a specific user (sub)population are hard to reach. This is often the case when trying to study and understand the issues or challenges immigrants face: their voice is often politically underrepresented, especially if they are outside mainstream

The associate editor coordinating the review of this manuscript and approving it for publication was Yichuan Jiang<sup>1</sup>.

society [3] or systemically excluded from national or international policy-making forums [4].

Ireland has recently grown into a multicultural society [5] with an increase in migration of foreign working professionals, international students, asylum seekers and refugees. Due to its welcoming policies confirming diversity and inclusion, Ireland has embraced these communities as their own [6]. Correspondingly, there has also been a significant growth in online activism on popular social networks, such as Twitter [7]. Hence, social media should be a valuable tool in trying to understand the non-native experience in countries like Ireland. This is our objective in this work: to try and ascertain a methodology that can be used to isolate (and later summarise) the voice of the non-native populations of Ireland to support more inclusive policy decisions.

In any society where multicultural identities are present, there remains a boundary between natives and migrants in

various aspects. More specifically, over the years, in cases of increased migration to European Union countries, there remains a certain level of scepticism on policies proposed for integrating migrants into society [8]. The general consensus, however, is that the policies need to be more effective in achieving their goals. The general problem with this consensus is that there has been limited quantitative analysis of migrants' lived experiences. One step towards addressing this is a better means of data collection and automated tools to assemble anonymous aggregated data that shed some light on any lived experiences. Nevertheless, this is not so straightforward.

Ironically, the over-representation of English on social media platforms becomes problematic.<sup>1</sup> In countries where English is not the dominant language, proficiency in the target language can be used as an indicator (e.g. [9]). This is not necessarily the case in English simply because its use is so widespread. Instead, researchers often focus on the meta-data of user accounts to help classify specific sociodemographic properties of users, for example, geo-tagged location data [10], [11] or other aspects of the user profile [12]. However, the main issue with this type of approach is that user metadata is often unreliable (due to issues such as self-representation where users present a socially stylised view of themselves [13], which can sufficiently distort meta-data potentially biasing such approaches) or large amounts of user data are needed to perform classification, which can be time-consuming to gather at scale (i.e. many users to process).

In this paper, we illustrate that telltale linguistic signs can be used to differentiate natives from immigrants in English-speaking countries such as Ireland. We hypothesise that how users express themselves online (i.e., how they use language rather than what they say) can be used to differentiate native Irish from non-natives. We also illustrate that this is achievable without the need for complex machine learning methods, large samples of data, or even user profile metadata. The latter (large data samples or user profile data) rather make models more accurate.

We focus on the case study of the Twitter landscape in Ireland, specifically examining the perspectives of migrant and native populations. We introduce a methodology that can accurately isolate the voice of Irish migrants using as few as 25 English tweets. We then use sentiment analysis, BERT-based topic modeling, and ChatGPT, to segment and summarize diverse perspectives in a way that policymakers can easily comprehend. Our methodology empowers policymakers to better comprehend multicultural communities and utilize this understanding to inform their decision-making.

The contributions of this paper can be summarised as follows:

- **Methodology for isolating the voice of non-native populations:** We propose a methodology that effectively

isolates the voice of Irish migrants on social media platforms, specifically focusing on Twitter. By analysing linguistic patterns in English tweets, we demonstrate that it is possible to differentiate between native Irish users and non-native immigrants without relying on complex machine learning methods or extensive user profile metadata.

- **Accurate classification with minimal data:** Our approach showcases that as few as 25 English tweets are sufficient for the accurate classification of users as natives or migrants. This highlights the potential for utilising limited data to gain valuable insights into the perspectives and experiences of immigrant populations, enabling a more comprehensive understanding of multicultural communities.
- **Sentiment analysis and topic modelling for diverse perspective summarisation:** We employ sentiment analysis, BERT-based topic modelling, and ChatGPT to segment and summarise diverse perspectives within the Twitter landscape of Ireland. This allows policymakers to easily comprehend the sentiments and concerns of both native and migrant populations, facilitating more inclusive policy decision-making.
- **Empowering policymakers with actionable insights:** Our research empowers policymakers to gain a deeper understanding of multicultural communities and their experiences in Ireland. By providing policymakers with accessible and interpretable summaries of diverse perspectives, our work supports the formulation of inclusive policies that address the specific needs and challenges faced by migrants.

Overall, our paper contributes to the field of inclusive policy-making by offering a practical and efficient NLP pipeline, InÉire, which supports inclusive policy-making concerning migrants in Ireland. Through our methodology and findings, we aim to bridge the gap between policymakers and migrant communities, fostering a more inclusive and informed decision-making process.

This paper is structured as follows. In Section II, we present and discuss key aspects of related work to situate our work in the literature. In Section III-B, we discuss (briefly) some key ethical considerations of this work that are necessary given its research objectives and context.

We then present our methodology in Section III. In Section III-A, we discuss (briefly) the dataset used in this paper, which we presented in [12]. In Section III-C, we present our approach to linguistically classifying Twitter users with a view to minimising the amount of user-generated content needed, and using only simple (yet interpretable) machine learning methods. In Section III-D, we describe our method of data segmentation via topic modeling; Section III-E explains how we leverage human-in-the-loop techniques to improve our methodology; Section III-F discusses the use of a summarization tool to generate human-understandable reports; and in Section III-G we describe our data preparation pipeline, which is the first and critical step in

<sup>1</sup>The proportion of content in English is extremely high. See: <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/> – last accessed 7th July 2023.

deploying data mining techniques in NLP tasks. We discuss how our data preparation pipeline is tailored to the specific task in each step.

To further illustrate the utility of our data processing pipeline, we conduct a case study on Irish social media content in Section V, which distills social media content into actionable insights for policymakers. Finally, in Section VI, we conclude the paper with an outlook to future work.

## II. RELATED WORK

Whilst concerns over excluding valuable insights (mainly when the target sub-population lacks sufficient internet connectivity [14]) exist, many studies have shown that this concern is rapidly waning. In the context of marginalised groups (which immigrants often fall into), social technologies (which include social media) are often essential [15], [16], [17], acting as a “safe place” due to aspects of anonymity and control over the intended audience [18]. The general availability of social technologies has also been high among marginalised groups for some time now [18], [19].

### A. SOCIAL MEDIA AS A LENS ON SOCIETY

In studies that leverage social media, we can easily distinguish studies that investigate more large-scale macro-level events, such as political campaigning (e.g. [20]), riots and civil unrest (e.g. [21]), event detection (e.g. [22], [23]), and large-scale news events like COVID-19 (e.g. [24]). Here, researchers often use social media as a lens to understand the perceptions and viewpoints of a subset of society or seek to identify key events that occur during the period(s) of observation. In these cases, it is generally easy to think about how to access and curate a meaningful and large corpus of text content. This is, however, not quite the case for studying immigrants: they are, as a hard-to-reach and arguably marginalised group, hard(er) to locate and study with social media. In fact, when it comes to marginalised groups, there is a general observation that researchers lack robust methodological guidelines [25].

There has been a lot of work (e.g. [26], [27], [28], [29]) studying hate, toxicity, and cyberbullying, which is often directed towards immigrants and marginalised groups via social media. Though useful from a methods perspective, these studies often have a strong content bias towards content stemming from North America and the United Kingdom [30]. [31] outlines how social media can be used to specifically study marginalised groups and discusses a number of the associated challenges. The key here is to ensure that the corpus is representative and not just an echo-chamber of a few (obvious) selected topics.

As discussed in [12] and [31] there are many strategies for curating a meaningful dataset for studying marginalised and immigrant communities. We reuse the dataset collected in [12] for this work to enable an illustrated comparison between an approach that utilises content and metadata from one that leverages linguistic properties of text. In [12],

we adopted the approach of assembling a manually curated set of users (as proposed in [32]) with the “activism” of a user being the key focus. This enables the study of a diverse range of topics emanating from marginalised communities versus privileged ones, with previous such studies limiting their study around an event, hashtag or group of similar users [33], [34], [35]. However, there is still the challenge of accessing new users. In such cases, we do not necessarily wish to extract a large portion of users’ Twitter history for classification purposes. Herein lies the main objective of this work: to classify a user with as little of their user-generated content as possible.

### B. KEY WORK IN COMPUTATIONAL LINGUISTICS

The objectives of this article mean that our work is related to studies on Native Language Identification (NLI). The goal of NLI is to detect the native language of an author, given a piece of her/his writing in a foreign language. Most of the research on NLI has been performed on identifying the native language of non-native English writers [36]. Most of the works on NLI use classification algorithms along with a range of features.

In order to identify (or classify) a social media user as a native or non-native, we need to cast the problem as a binary classification problem. There has been significant work on text classification with social media data. [37], [38] provide a good general introduction to the general methods of social media analytics, and our work maps well to their general implementation architecture. Twitter has been the platform of choice for many years, and the findings of [39] illustrate this to still be the case<sup>2</sup> with a major emphasis on binary classification tasks and simple off-the-shelf machine learning techniques (i.e. Decision Trees, Support Vector Machines, and Naïve Bayes, we also know from [40] that simple models are often sufficient for a large variety of machine learning tasks). While there is an increased interest in deep learning, it is both more demanding of training data and can obscure the learning process which limits interpretability.

When looking at methodological guidance, [41] provides a general review of text-based analysis of social media data, specifically with an in-depth review of Twitter analyses. Yet, our work suffers from the challenge of a lack of labelled data; a problem shared by hate speech detection (e.g. [21], [26], [27], [42], [43]), extremist content (e.g. [44]), and other forms of distress-based classification (e.g. [45], [46]) all of which can help inform our approach. Yet, we also have the challenge of small user samples, a challenge shared with the task of dialect classification (e.g. [47]). Hence, we build on our previous work (and data curated) in [12].

Returning to NLI, most approaches have been trained and evaluated using the TOEFL11 dataset [48], which has been the standard dataset for this task. TOEFL11 contains about 13,000 English essays written by English learners with 11 different first languages for the TOEFL test. Recent approaches

<sup>2</sup>However, we note the recent changes to the API for academic researchers may change this in time.

to NLI have focused on social media content to detect the native language of users. Our work falls into this line of research. Since the proficiency level of English authors on social media sites is much higher than that of the learner datasets (such as TOEFL11) [49], the NLI task using social media is much more challenging than in the learner datasets.

Reference [50] built a dataset containing English tweets written by native speakers from 12 other languages besides English. They use this corpus to train a model to predict the native language of users based on their English tweets. Reference [51] attempts to predict the native language of Reddit users using both linguistically-motivated features and the characteristics of the social media outlet. Reference [52] uses BERT (Bidirectional Encoder Representations from Transformers) to detect the native language of the Reddit authors. Reference [50] predicts the native language of users based on about 173 tweets on average. References [51] and [52] try to infer the native language of a user based on 100 Reddit posts. Here, it is also useful to try and capture some discussion on data quantity. In this work, we predict if a user is native (or close to it, for example, if they have lived in an English-speaking country for a long time) of English or not, based solely on 25 tweets and using only linguistic properties of the tweet content. We do so based on the observations of [13], who noted that small amounts of social media content (in this case, 50 words) are sufficient to linguistically profile users even for quite complex machine learning tasks (personality profiling in [13]).

In terms of basic text-mining methodology for classification, our approach is well grounded in the literature following [53]: pre-processing and feature extraction using LIWC [54], [55], followed by a simple machine learning model (based on [40]). We note that our use of LIWC allows us to focus not on what users say, but rather how they express themselves, this follows on from our previous work [13], [20] where the use of LIWC as a method of feature engineering permitted complex research questions, with simple machine learning methods. It has also been shown that combining machine learning methods with LIWC improves performance [56]. LIWC identifies portions of text that belong to specific linguistic categories of text used in the tweet. This means our classifier leverages a feature set that sits at the intersection of part-of-speech tagging (e.g. nouns, pronouns, verbs etc.) affect analysis (forms of emotion in the text, beyond just positive and negative) and other linguistically-orientated traits in the text that can be indicative of different characteristics of users (e.g. average word length, use of singular vs. plural pronouns, presence of non-fluencies, use of function words etc.) [57], [58], [59] that have been used to build machine learning models complex computational social science problems.

### III. INÉIRE DATA PROCESSING METHODOLOGY

In this section, we present a data pipeline that can support policymakers in discovering and analysing the topics that migrants and natives in Ireland discuss on social media, which

has the potential to inform the policies that affect migrants and other marginalised populations in Ireland. Figure 1 shows an overview of the proposed approach capturing four key steps: 1) classification of social media users with simple and explainable supervised machine learning; 2) content segmentation via sentiment polarity to differentiate between positive and negative content; 3) clustering of content into topics to enable further segmentation of the data into positive vs negative topics of discussion; employing a human in the loop to review topics generated (i.e. remove meaningless topics or topics of no relevance); and finally, 4) we leverage ChatGPT as a summarisation tool to distil each topic into a brief summary of the main themes. In the subsequent subsections, we discuss each step of our methodology in detail, but first, we present the dataset used throughout this work.

#### A. DATASET

The Twitter dataset [12] used in this study comprises both native (or near to it, for example, if the user has lived in Ireland for a long time) and non-native to Ireland Twitter users.<sup>3</sup> To curate the dataset, we employed a two-fold approach. Firstly, we utilised the dataset provided by Younus et al. [12], which comprises tweets from native and non-native users in Ireland. This dataset allowed us to gather a comprehensive collection of tweets related to Irish migrant and native activism. Secondly, we curated a list of Twitter activists from both migrant and native communities in Ireland. The curation process was performed by one of the authors familiar with Ireland's social justice landscape and capable of distinguishing between native and migrant users. Since Ireland is an English-speaking country, we consider Irish users as native English speakers, and the migrants are regarded as non-native English speakers.<sup>4</sup> This yielded a dataset containing (mostly) English tweets,<sup>5</sup> arranged into native and migrant tweets. The following defines the inclusion/differentiation curation strategy of natives and migrants on the Twitter dataset:

- Checking the surnames of the users against Irish surnames.<sup>6</sup>
- Reading the biography field of the user and checking for Irish terms in addition to flags of various countries.<sup>7</sup>
- Reading the last 20-100 tweets of a user to see whether there is any explicit mention of belonging to any country.

This curation process resulted in a dataset consisting of tweets classified as either native or migrant tweets. We aimed

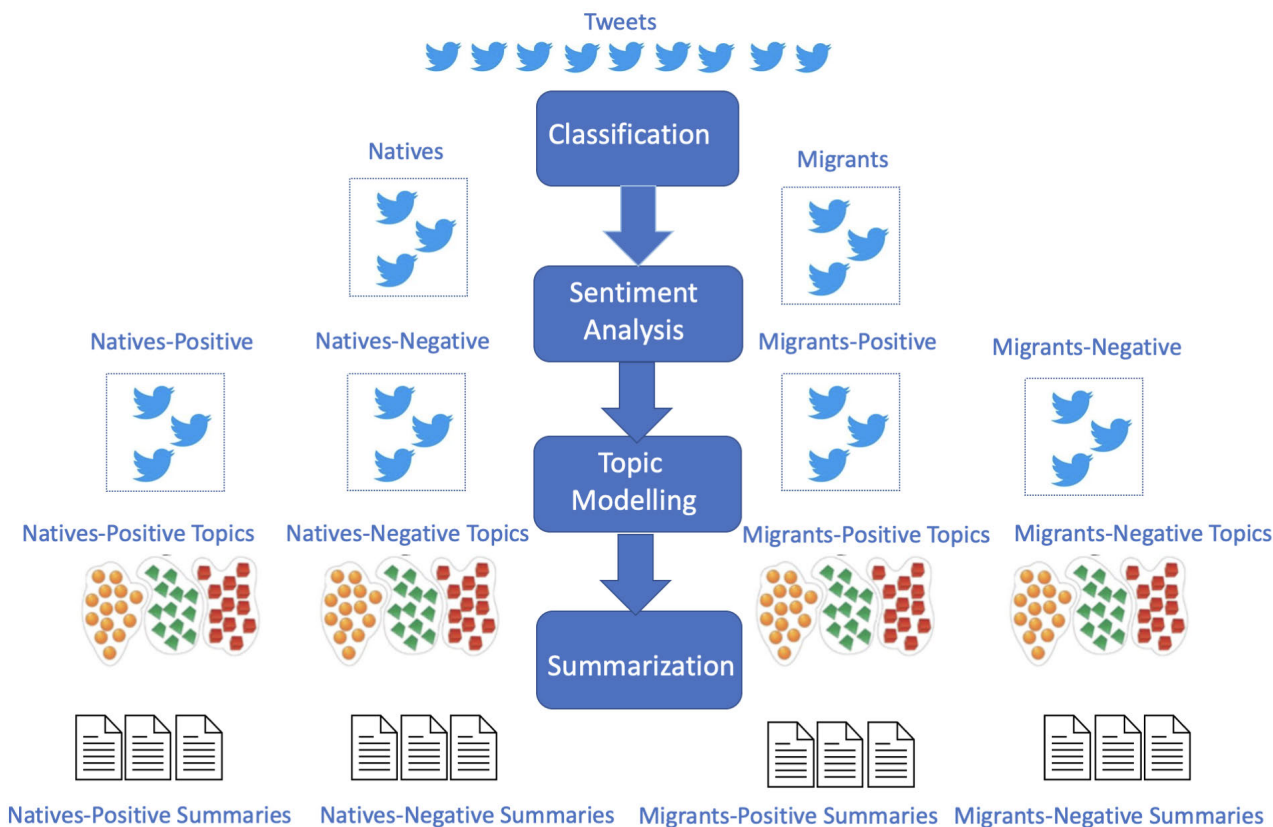
<sup>3</sup>Note that an anonymised version of this dataset has been released for public download at <https://github.com/InEire>. The companion repository for this article can also be found using the same link.

<sup>4</sup>Note that this could be problematic in cases where a migrant originates from an English-speaking country, such as the UK.

<sup>5</sup>We automatically detect and filter out non-English tweets and retweets using standard python-based tools retaining original English user-generated content, i.e. retweets are also filtered out of the data as these are arguably not representative of the (socio)linguistic profile of the user.

<sup>6</sup>Performed using <https://www.duchas.ie/en/nom>

<sup>7</sup>It has been observed that migrants within Ireland usually insert a flag representing Ireland and their country of origin.



**FIGURE 1.** High-level overview of the approach showing the four key steps: 1) classification of tweets written by natives or migrants; 2) sentiment analysis of tweets; 3) topic modelling; 4) summarization.

to maintain a balanced representation of 66 native and 66 migrant activists within the dataset. Although the sample size is relatively small, it provides a fair representation of the Irish Twitter landscape, considering the country’s size. The dataset encompasses mostly English tweets, with non-English tweets and retweets filtered out to retain original English user-generated content. Additionally, we focused on non-native speakers who have resided in Ireland, ensuring a higher level of English proficiency among the migrant user subset.

To ensure data privacy, we employed standard practices and adhered to ethical guidelines when collecting and processing the data. As an academic research project, we obtained access to the Twitter API for research purposes and extracted the last 3000 tweets from each user in the curated list, resulting in approximately 300,000 tweets from both native and migrant users.

Following standard supervised machine learning practices, we use a holdout strategy to generate three independent training, validation (or tuning) and testing subsets of the data containing 400,000, 100,000, and 100,000 tweets, respectively. Note that this sampling is at the user level instead of the Tweet level, i.e. a user (and their collected Tweets) is placed into one of three sets, i.e., train, validation, or test set. We note here the findings of [60], i.e., that resampling our

dataset could yield structurally different results. Nevertheless, as our objective is to (partially) explore the robustness of the approach, it does not induce significant concerns.

**B. ETHICAL AND BIAS DELIBERATION**

The presence of specific ethical concerns and/or other biases when using social media data has been discussed at length in the literature (for example [14], [38], [61], [62]). The source of many ethical concerns around social media often stems from the ability to gather and therefore analyse large quantities of data [62]. Yet, taken out of context, this data may lose its meaning [63] or be otherwise warped beyond its original communicative intent. This is key, as social media is often leveraged for information of the now, but when used historically (as in this paper), it can be difficult to access representative data [64]. Many studies (e.g. [60], [63], [64], [65]) have discussed this at length. Essentially, the key take-aways here are that data may be missing, and that changing the sample of data used (e.g. due to it being missing) can fundamentally change the findings and the interpretation of the data.

A severe grey area in social media research is the issue of data availability – generally, it is available and easy to access, which gave rise to many controversial studies (like the Cambridge Analytica Scandal). Many problems have

stemmed from data being publicly available (on social media platforms) and freely “browsable” but where researchers had permission from the platform though not the users themselves, to process the data. In this sense, there may be ethical tensions in the use of social media content as a means to study potentially vulnerable populations such as migrants. It may also be difficult to predict potential harms from the analyses [31], [63]. A key area of ethical debate in social media research is informed consent. Unlike other research methodologies, social media users are themselves not really “participants” in a traditional sense, and are often unaware that their publicly available data is being used. As discussed by [63] and [66] it is rather important that data that would otherwise have been private remains so. This affects the extent to which informed consent of users should be sought in order to comply with data protection laws. This does not mean that data in the public sphere is fair game because a user has “agreed” to a set of terms and conditions from the platform. Instead, it is key to ensure anonymity and minimise any potential risk(s) of harm [67], which, in this case, informs the extent to which data can (or even should) be made available following on from any analysis. Thus, in our context, care is needed to ensure that users cannot be spotlighted or identified as a result of this work. With these concerns in mind, we are acutely aware that high levels of caution are needed in studies such as this. We cannot be too invasive into the (sub)population(s) studied, yet we still need to garner meaningful results, albeit at an aggregate level. We also note, at this point, that this work has been reviewed by our institutional ethics committees.<sup>8</sup>

Related to the ethical deliberation of this work are considerations of potential biases. For example, self-representation – the promotion of an often stylised “voice” of the individual catered for the expected audience (see [13] for an overview). Self-representation is akin to a common method bias for social media researchers, and there are not many easy-to-use tools to combat it. Whilst not specific to this work, it would be naïve to not expect some amount of self-representation or social censoring in the sense of not voicing specific concerns publicly. Thus, while we can be confident that there is some value in exploring the migrant voice via social media, it should not be considered as a complete representation of this voice. Similarly, purely content-based analyses need to be considerate of such challenges.

We must also be aware that migration is often a hotly discussed topic. Specifically for studies that seek to leverage social media as a source of data for scientific study, there is an abundance of additional challenges related to the quality and/or veracity of information communicated, i.e. misinformation. This is not necessarily always malicious, however, there are many cases where it is, as shown by [68], [69], [70], [71], and [72]. Taking all these factors into account,

<sup>8</sup>Specifically, under REC-20-224 (TU Dublin) and LS-E-21-257-Caton (UCD).

we look for telltale linguistic signs (using LIWC as a method of feature engineering) that focus on how a social media user writes as opposed to what they actually say when classifying users into specific categories.

### C. USER CLASSIFICATION

We aim to select a set of  $k$  tweets from a user and determine whether this user is a native of Ireland or a migrant. To define the problem formally, assume that we have a set of  $k$  tweets from a user as Equation 1, where  $Tweet_{i,j}$  is the  $j^{th}$  tweet of user  $i$ .

$$TweetSet_i = [Tweet_{i,1}, Tweet_{i,2}, \dots, Tweet_{i,k}] \quad (1)$$

Each  $Tweet_{i,j}$  contains a set of tokens as shown in Equation 2, where  $t_{i,j,m}$  shows the  $m^{th}$  token of the  $j^{th}$  tweet of user  $i$ , and  $l$  is the number of tokens in  $Tweet_{i,j}$ .

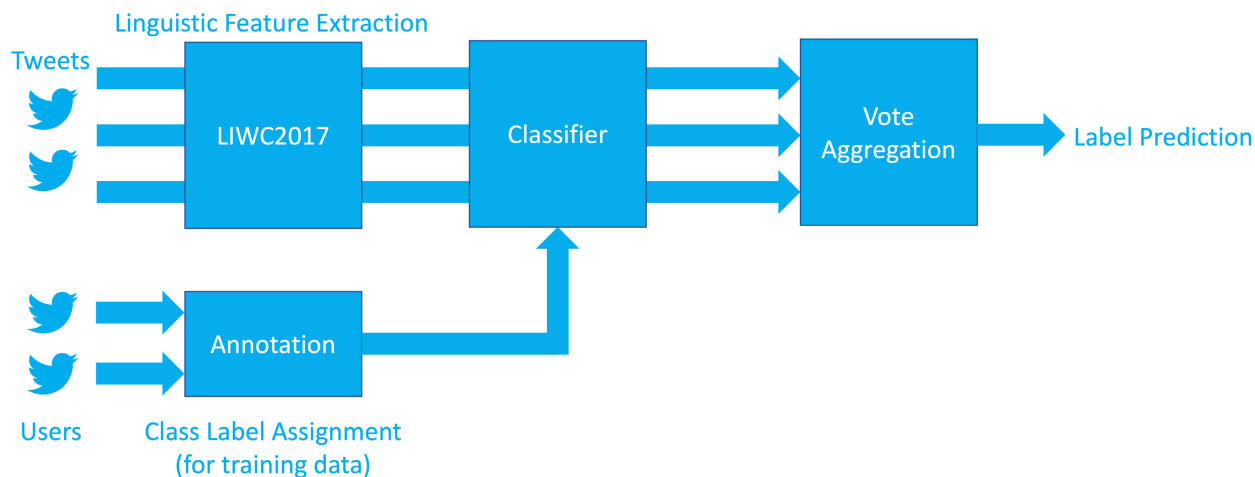
$$Tweet_{i,j} = [t_{i,j,1}, t_{i,j,2}, \dots, t_{i,j,l}] \quad (2)$$

We aim to map each  $TweetSet_i$  to a  $label \in \{0, 1\}$ .  $label = 1$  indicates that the user is native, and  $label = 0$  indicates that the user is a migrant. We employ supervised machine learning to predict the label for each user and solve the stated problem. Supervised machine learning algorithms require a reasonable amount of annotated data. For our task, the annotated data are in the form  $(Tweet_{i,j}, label_i)$  where  $(Tweet_{i,j})$  is  $j^{th}$  tweet of user  $i$ , and  $label_i \in \{0, 1\}$  shows whether the user  $i$  is native or migrant. We built a dataset containing 600,000 tweets from natives and migrants from Ireland Twitter users.

Our proposed method consists of three main components: (1) Feature Extraction: extracting linguistic detail from users’ tweets; (2) Classification: the training and a (simple) machine learning model at the tweet level, i.e., does a specific tweet appear to belong to a native or migrant; and (3) Majority Voting: use each of the  $k$  tweets from one user cast one vote, i.e. native or migrant as a decision aggregation function. This is summarised in Figure 2.

We use Linguistic Inquiry and Word Count (LIWC) to extract features from the tweet texts. LIWC is a text analysis tool that extracts features from natural language text [54], [59]. The features extracted by LIWC are summary language features (analytical thinking, clout, authenticity, and emotional tone), standard linguistic features (e.g. frequency of pronouns, articles, function words, etc. in the text), word categories tapping psychological constructs (e.g., affect, cognition, biological processes, etc.), personal concern categories (e.g., work, home, leisure activities), informal language markers (assent, fillers, netspeak, swear words) and punctuation categories (periods, commas, etc.) [55]. They capture more of how a user communicates rather than what their communication act or message is.

We use LIWC because we aim to model the writing style of users, and LIWC extracts various linguistic features suitable for this purpose. We used LIWC 2017 and extracted 93 features from each tweet’s text in our dataset. The feature extraction component receives  $Tweet_{i,j} = [t_{i,j,1}, t_{i,j,2}, \dots, t_{i,j,l}]$  instances in our dataset as input, and produces a feature vector



**FIGURE 2.** Three key components in classification step: 1) linguistic feature extraction with LIWC; 2) classification of tweets; 3) voting aggregation to produce a user label prediction.

as in Equation 3 as output. In Equation 3  $f_{i,j,m}$  indicates the  $m^{th}$  feature produced by LIWC, for  $j^{th}$  tweet of user  $i$ .

$$FeatureVector(Tweet_{i,j}) = [f_{i,j,1}, f_{i,j,2}, \dots, f_{i,j,93}] \quad (3)$$

We cast the problem as a supervised binary classification task and used a decision tree, an interpretable method, for the classification. We opt for a decision tree because it is relatively simple, fast, and highly interpretable. The latter is specifically poignant as it will also reveal which LIWC categories (features of  $Tweet_{i,j}$ ) are most indicative. The decision tree receives  $FeatureVector(Tweet_{i,j})$  as input and predicts  $label_{i,j}$  to suggest whether  $Tweet_{i,j}$  is written by a native or migrant to Ireland. We implemented the decision tree using Python’s sklearn library [73]. It is also worth reiterating here that there is no significant effort to use complex machine learning methods; the classification task operates at the tweet level, and we sample  $k$  tweets from a user to generate a view of their linguistic profile. This reduces specific content biases that can be encoded: we do not look at content; and also use multiple tweets to profile a user.

As mentioned earlier, we aim to get a set of  $k$  tweets from a user and determine whether this user is native or not. Since the non-native users might have some tweets written exactly in the same way as natives (for example, typical tweets such as: “Thanks!”, “Great!”, “Congrats”, etc.), we use majority voting on the labels predicted for the  $k$  tweets of each user. This affords a degree of robustness against “outlier” or unusual tweets a user may have, and similarly, very short or uninformative (from the perspective of classification) tweets.

To this end, we keep the original order of the tweets and segment the dataset into chunks of  $k$  tweets. Each chunk contains the  $k$  consecutive tweets written by the same user, as in Equation 1. Then we predict the label of each  $Tweet_{i,j}$  using the decision tree described earlier. Finally, we use majority voting to aggregate the labels of the tweets and predict the final overall label for the user. If most of the  $k$  tweets in

$TweetSet_i$  had  $label = 1$ , we consider the final label as native. Otherwise, we consider it as migrant. We concede that there is a potentially important decision (or hyperparameter) here: the extent of the majority. However, in such cases, it may, in fact, be easier (initially) to just increase the size of  $k$ . However, this is a valid concern (the size of  $k$  compared to the extent of the majority) that we can explore empirically.

#### D. DATA SEGMENTATION VIA SENTIMENT AND TOPIC MODELING

After classifying Twitter users into native and migrant groups, we apply sentiment analysis to their tweet corpora (using Vader [74]). This enables us to differentiate online discourse into different polarities aligned to different (sub)populations and obtain four tweet sets: (1) Natives-Positive, (2) Natives-Negative, (3) Migrants-Positive, and (4) Migrants-Negative. Upon these tweet corpora, we perform topic modelling to find potentially relevant topics that could be used to inform policy makers.

There are various techniques to perform topic modelling on social media content, such as tweets. The most popular topic modelling techniques are Latent Dirichlet allocation (LDA) [75] and its variations. LDA is a statistical approach to topic modelling. In LDA, each tweet is considered a probabilistic mixture of topics, and each topic is considered a probabilistic mixture of words. LDA uses the bag-of-words representation of the tweets to find the corresponding topic. BERTopic [76] is a recent approach to topic modelling which leverages BERT embedding and class-based TF-IDFs of the tweets to generate coherent topics. In our preliminary experiments, we examined both LDA (and its variations) and BERTopic. However, BERTopic produced more coherent and stable topics motivating its use in this study. Applying BERTopic generates 15 topics for each of our four tweet sets and also identifies the tweets that correspondingly encompass each topic. Thus, we use it more as a means to cluster tweets

according to their content to facilitate the exploration of topical discourse among the two user populations. For the next steps and further analysing, we choose the most sensible topics among the 15 topics for each tweet set.

### E. HUMAN IN THE LOOP

BERTopic is a powerful unsupervised topic modeling tool that has been shown to outperform other state-of-the-art techniques. However, relying solely on automatic methods can lead to including irrelevant topics, which can result in inaccurate and incomplete results.

To address this issue, we employ human-in-the-loop (HitL) to review and filter the topics identified by BERTopic. In this step, we involve domain experts who have a deep understanding of the immigration-related issues and can provide valuable insights into which topics are relevant for policymakers. The experts review the topics and investigate the relationship among them, helping to identify the relevant topics that policymakers are most likely to be interested in.

After identifying the relevant topics, we feed the tweets that correspond to these topics into a summarizing tool to generate summaries of the discussions related to the identified topics. These summaries provide a concise and comprehensive overview of the key points discussed in the tweets, allowing policymakers to quickly gain insights into the related issues among migrants and natives populations. Whilst for this work, we use one of the authors as a domain “expert” in a real deployment of the methodology we could replace this with a committee of immigration experts to review and select topics for further processing and consideration in the pipeline.

### F. SUMMARIZATION

As mentioned earlier, the ultimate goal of this study is to enable policymakers and other related agencies to further their understanding of native and migrant communities in Ireland. Topic modelling extracts the topics discussed by natives and migrants on social media. At the tweet level, topic modelling determines the topic discussed in each tweet. We aggregate the tweets based on their topics and use the aggregated tweets within a topic to establish what users say about a specific topic.

To produce reports that are comprehensible to policymakers, we use a summarisation tool and summarise the aggregated tweets. Specifically, we employ ChatGPT. ChatGPT is a large pre-trained language model developed by OpenAI which is able to generate human-like responses in natural language conversations. We provide the tweets in each topic as input to the ChatGPT prompt and ask it to summarise the tweets. Given that ChatGPT is restricted by input length, we divide the tweets into segments, each containing approximately 70 tweets, and prompt ChatGPT to summarize each segment. We note that as the ChatGPT (and related tools) API develops, that this sample size can be increased. Similarly, paid variants of the API can also support larger numbers of tweets for summarisation. However, our experience (even

**TABLE 1. Tools used for each of the pre-processing steps.**

Pre-processing Step	Tool
Removal of retweets, Removal of mentions	Python
Tokenization, Removal of stop words	NLTK Platform
Removal of URLs, Removal of hashtags	Preprocessor package
Removal of emojis and smiles	Preprocessor package

with arguably small samples of tweets) has been positive as we will discuss in section V.

### G. DATA PREPARATION

Data preparation is the first and critical step in deploying data mining techniques in NLP tasks. It has a significant impact on the success of the data mining process. As outlined in this section, our data pipeline consists of four main steps: Classification, Sentiment Analysis, Topic Modelling, and Summarisation. Based on the specific task in each step, we use a different pipeline for data preparation as illustrated in Figure 3.

First, we excluded retweets, as they are not authored by users. Then we perform word tokenisation and split the text of the tweets into tokens. Finally, we remove mentions from the tweet texts (as they typically not going to be useful later in the pipeline).

In the Classification step, we aim to select a set of tweets from a user and determine whether this user is a native of Ireland or a migrant. In this step, we utilise the linguistic patterns in the tweet text to separate natives’ and migrants’ tweets. Hence, we do not perform any further pre-processing tasks to keep the writing style of the users.

For the Sentiment Analysis step, we remove URLs and Hashtags from tweets since they do not have a direct impact on the sentiment of the tweets. For the topic modeling step, we remove URLs, emojis, smilies, and stop words. We keep the hashtags because they assist topic modeling in detecting the correct topics of the tweets (they are after all used by users to tag the content of their tweet). We remove stop words because they are frequent in the text, and topic modelling techniques may misdiagnose them as the main topics of the text. For the summarisation step, we remove URLs, emojis and smilies. We keep hashtags and stopwords because summarisation techniques require them to comprehend the meaning of the text and summarise it. Figure 3 shows the data preparation pipeline. Table 1 shows the tools we used for performing each of the pre-processing steps.

## IV. CLASSIFICATION PERFORMANCE

We perform a binary classification task that predicts a label (native, non-native) for each tweet. We use four popular evaluation metrics over tweets: Accuracy, Precision, Recall, and  $F_1$ . The classifier is trained on the training set and evaluated on both the validation and test sets using a standard holdout strategy, employing a train-validation-test split of 400k/100k/100k tweets. The final label for each user is determined based on  $k$  consecutive tweets, and in our experiments,



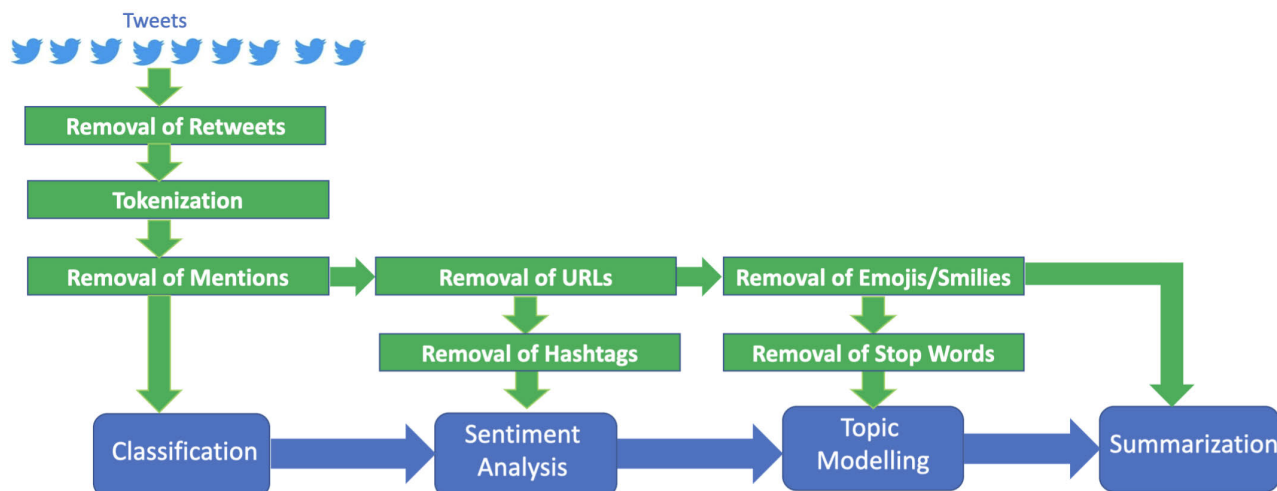


FIGURE 3. Data pre-processing.

we investigate the effect of the value of  $k$  (from 1 to 200) on the performance of the classifier. Ideally, we aim for a small value of  $k$  to enable confident classification of new users to be confident without accessing a large portion of their Twitter history. Figures 4a, 4b, 4c and 4d show the impact of  $k$  on the accuracy, precision, recall, and  $F_1$  of the classifier on the validation set.

It can be seen that increasing the value of  $k$  increases the classifier’s performance for all evaluation metrics (accuracy, precision, recall, and  $F_1$ ). All figures show that having more tweets from a user improves classifier performance, which is not surprising. Yet, increasing the value of  $k$  from 1 to 25 has maximal impact; there is arguably only marginal improvement above  $k = 25$ , and certainly slight improvement above  $k = 50$ . This yields an interesting finding that with about 25 tweets, we can predict reasonably well (an accuracy of about 80%) whether a user is a native of Ireland or not.

Figure 4b shows that for all values of  $k$ , the precision of the classifier for the “native” class is higher than its precision in the “non-native” class. For  $k > 25$  the precision of the native class exceeds 90%. This shows that the classifier can detect “native” users with a high confidence. Similarly, 4c shows that for all values of  $k$ , the recall of the classifier in the “non-native” class is higher than the “native” class. For  $k > 25$  the recall of the “non-native” class exceeds 90%. This means that the classifier can find “non-natives” very accurately. Figure 4d shows that the performance of the classifier in terms of  $F_1$  score on the “non-native” class is better than the “native” class illustrating that the balance between precision and recall is better for this class. This is actually key for our purposes, specifically with the objective of being able to better understand the migrant view(s) of Ireland. Choosing  $k = 25$  and evaluating the model on the test set yields the results illustrated in Table 2: the model performs well on the unseen test set.

In contrast to the results presented in [12], which employed a larger dataset consisting of Twitter data and associated

TABLE 2. The performance of the model against method [12], on the test set.

Proposed Method			
Accuracy	80%		
Natives	Precision	Recall	$F_1$
	87%	71%	78%
Non-natives	Precision	Recall	$F_1$
	75%	90%	82%
[12]			
Accuracy	85-87%		
Precision	78-80%		

metadata, this study adopts a smaller sample size. Despite this disparity, our approach achieves a slightly lower accuracy rate (80% vs. 85-87%), but achieves a better native precision and migrant recall. An interesting observation is that we achieve these improved metrics using significantly less Twitter data and without relying on metadata. This suggests the efficacy of our approach in isolating the migrant voice using an interpretable NLP technique.

We emphasise that our study focuses specifically on the task of isolating the Irish migrant voice in the Twitter landscape, utilising a distinct methodology and emphasising interpretability. The work by Younus et al. [12], which generated the dataset, differs in terms of its objectives and the inclusion of metadata. While [12] provides valuable insights into a broader set of research questions, our work aims to address the specific challenge of supporting inclusive policy-making concerning migrants in Ireland. By leveraging a simpler and more interpretable approach, we aim to distil differentiated views in a consumable manner for policymakers.

As mentioned in Section III-C we used all of the 93 features extracted by LIWC for the decision tree. LIWC produces linguistic, thematic and psychological features to describe the text. This allows us to also investigate the importance of each LIWC feature in distinguishing between native and non-native tweets. To this end, we inspect the feature importance of the produced model. The higher the value, the more

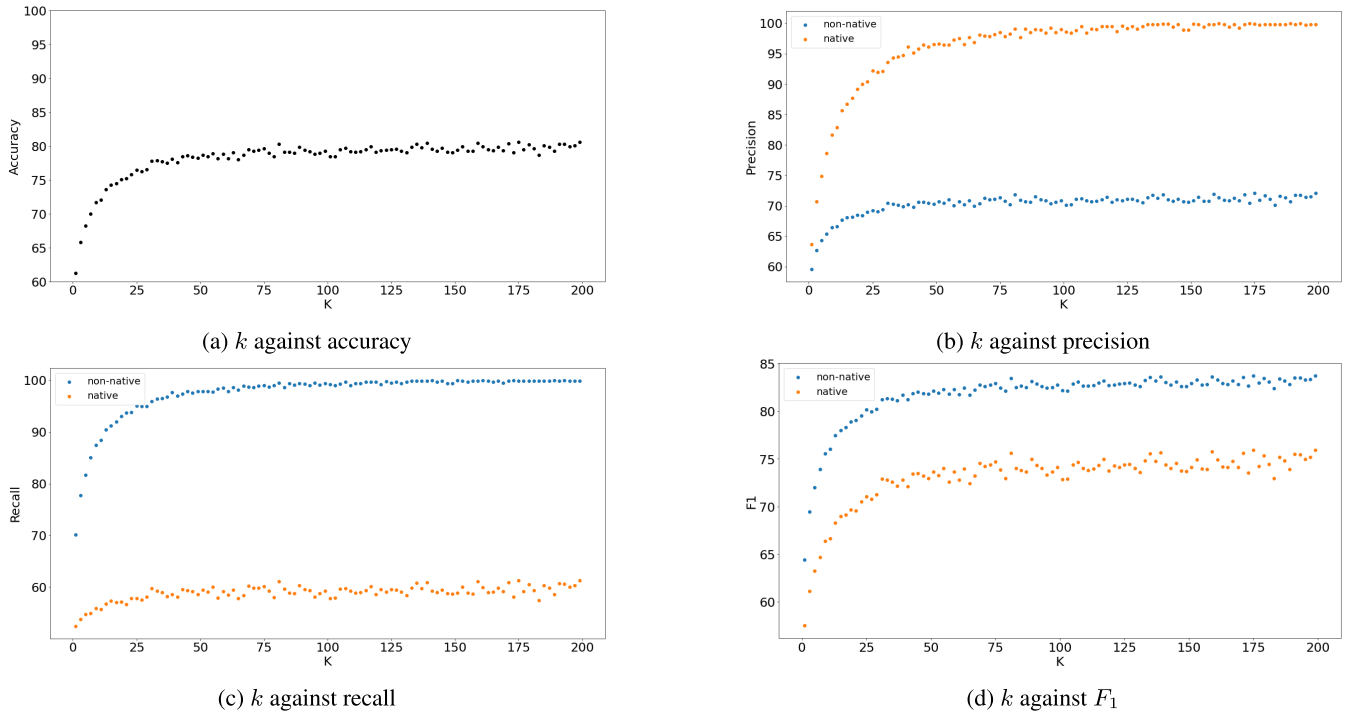


FIGURE 4. The impact of  $k$  on performance: a) accuracy; b) precision; c) recall; and d)  $F_1$ . Note: the y-axes do not start at 0.

TABLE 3. The LIWC feature’s importance in distinguishing between native and non-native english tweets.

Rank	LIWC feature	Importance
1	prep, OtherP, ppron	[0.03-0.04]
2	Dic, AllPunc, SixLtr, WPS, netspeak, Period, function	[0.02-0.03]
3	focuspresent, social, Analytic, relativ, verb, you, adj, article, drives,cogproc, Comma, time, work, number WC, pronoun, Apostro, adverb, Exclam, Clout, conj, space, Authentic, auxverb, Dash, ipron, power, affect, focuspast, informal	[0.01-0.02]
4	leisure, percept, hear, relig, Colon, quant	[0.009-0.01]
5	posemo, affiliation, i, reward see, cause, tentat, assent	[0.008-0.009]
6	certain, compare, insight,achieve motion, differ	[0.007-0.008]
7	negemo, Tone, bio, interrog discrep, focusfuture, health, negate	[0.006-0.007]
8	SemiC, anger,shehe	[0.005-0.006]
9	money, friend, Quote, home, risk, QMark	[0.004-0.005]
10	we, female, they, family, male, sad	[0.003-0.004]
11	nonflu, body, anx, feel, death, Parenth, ingest, sexual	[0.002-0.003]
12	filler	[0.001-0.002]
13	swear	[0-0.001]

important the feature in labelling the tweets. Table 3 shows the importance of different LIWC features.

The first 10 important features in Table 3, along with their description, are shown in Table 4. Interestingly, all of the

TABLE 4. The 10 most important LIWC features along with their description.

LIWC feature	Description
Prep	Prepositions (to, with, above)
OtherP	Other Punctuation
ppron	Personal Pronouns (I, them, her)
Dic	Dictionary Words
AllPunc	All Punctuations
SixLtr	Words with > 6 letters
WPS	Words Per Sentence
netspeak	netspeak (btw, lol, thx)
Period	period (.)
function	function words (it, to, no, very)

first 10 important features are standard linguistic features, informal language markers and punctuation categories. Features corresponding to psychological constructs and personal concern categories are not among the important features.

If we further inspect the 10 most important LIWC features, we see some key takeaways of interest. Natives usually write longer tweets, use terms more likely to be found in the LIWC dictionary, and use more function words and prepositions. In contrast, non-natives usually use more netspeak, and also have a higher usage of personal pronouns and punctuation.

This yields an interesting observation that distinguishing between native and non-native English speakers has been performed mostly based on the linguistic features of the tweet text, i.e. how a user uses language is highly correlated to their corresponding class label. This would also suggest that the model is generalisable in the sense that it is relatively content agnostic and also not dependent on potentially (socially)

stylised metadata as well as requiring only a small  $k \approx 25$  tweets.

## V. CASE STUDY ON IRISH SOCIAL MEDIA CONTENT

To illustrate the utility of the InÉire data processing pipeline, we conducted a small-scale case study using all components. The case study aims to highlight the practical implications for policymakers, i.e. distil a set of social media content (tweets) into actionable insights: topics, concerns, and issues discussed on social media platforms, in this case, Twitter. We note that the “findings” of this case study will be biased towards the sample of users in the test set and, therefore, not necessarily representative of Irish sentiment or discourse online. Thus, caution is needed in interpreting topics presented here as outcomes of this process. We leverage three parts of the InÉire methodology: topic extraction (V-A) combined with human-in-the-loop (V-B), topic filtering and summarisation (V-C).

### A. TOPIC MODELING AND EXTRACTION

We use BERTopic to extract the topics that migrants and natives discuss. BERTopic extracted 15 topics from each tweet set (Migrants-Positive, Migrants-Negative, Natives-Positive, Natives-Negative). Table 5 shows the most sensible topics extracted by BERTopic for each subcategory: positive and negative polarity tweets for the Migrants and Natives sub-populations. Note that this categorisation is determined in this instance by the classification model on only the test sample (III-C) and not by the ground truth label in an attempt to be as authentic an orchestration of the methodology as possible. Here, we define “sensible” topics as ones that are obviously meaningful in the presence of a little localised domain knowledge. For example, topics with disproportionately large numbers of tweets associated with them or with several “filler” or stop words in their generated title would not be considered “sensible”.<sup>9</sup>

### B. HUMAN IN THE LOOP

Following the extraction of topics from the segmented (by type and polarity) corpora, we review topics for content that is potentially thematically relevant. For example, in Table 5, positive topics #1 and #4 for the Migrants are clearly irrelevant to policymakers, yet topics #2 and #3 likely are (Seanad Éireann is the name of the Irish Senate). We can add a little further context of the socio-political climate in Ireland at the time that the data was collected to give meaning to the remaining topics that are “of interest”.

By socio-political climate, the war in Ukraine was underway; as such, there is no surprise to see content around this humanitarian crisis, and similarly, with an influx of refugees, content around the impact of this would also be expected (e.g. asylum seeking, deportation, etc.). Also, Ireland (especially the capital Dublin) is in the wake of a housing crisis

<sup>9</sup>Note that the topic title captures the most representative words in the generated topic.

(a lack of affordable housing for purchase and rent), which has been exacerbated by both rises in the costs of living (war in Ukraine, and the associated energy crisis and rise in inflation) and additional pressures due to an increased population in already densely populated major cities (more international students: relaxation of COVID restrictions, potential impact(s) of Brexit, and higher levels of immigrants and asylum seekers). Similarly, school teachers (both primary and secondary) are also currently in short supply. There is also debate around vaccines, especially COVID boosters that currently have a low uptake in the 35-49 age range. Finally, as a point of reference, the term “traveller” refers to an indigenous ethnocultural group in Ireland, often in receipt of stereotypical forms of racism.

### C. TOPIC SUMMARIZATION

Once topics have been determined as thematically relevant, we use word clouds to create a simple visualisation of the topics: word clouds capture the relative frequency of terms in the corpus; bigger means more frequent. To illustrate the differences between relevant, irrelevant and potentially relevant topics, we can compare the word clouds displayed in Figure 5. Here, it is already obvious that the football topic (Figure 5a.) is not relevant to policymakers, thus necessitating a human in the loop part of the process. While word clouds are useful for a quick sanity check of a topic, they do not really capture the nature of the discourse beyond the sentiment tag (positive or negative) and a set of frequent words. Thus, to provide more clarity and meaning for policymakers, we leverage ChatGPT as a more sophisticated content summarisation tool.

Table 6 shows the summaries generated by ChatGPT for the topics captured in Figure 5. As the number of Tweets corresponding to a topic can be quite high, we randomly sample ( $n = 150$ ) without replacement from the topic cluster. We note the findings of [60] in doing so, i.e. that the sample can dramatically affect the interpretation. Yet, our goal here is not to perfectly summarise the content but rather give policymakers an insight into social media discourse and specifically highlight or explain potential areas that their policy decisions should consider, specifically around differences between migrant and native populations. In other words, try to make them aware of what they are (potentially) not aware of in terms of potential new policy considerations.

### D. DETAILED INSIGHTS FROM THE INÉIRE INTERPRETABLE NLP PIPELINE

We can already see some key differences in Table 5. The positive topics are a collection of largely hedonistic and eudaimonic topics, with some politically focused ones as well. Yet, there are quite striking differences in both the positive and negative topics. In terms of positive topics, migrants’ focus lies on the day-to-day joys of life in Ireland, particularly evident from positive topics #1, #3 and #6 (from Table 5), whereas natives appreciated the creation of better cycling routes (natives’ topic #1 from Table 5) along with Ireland’s commitment towards Ukraine and its firm stance

**TABLE 5.** The generated topics by BERTopic for each tweet set, ordered by size. We also show the number of tweets for each topic in brackets. To give context, there were about 30'000 tweets in this portion of the dataset.

Topic Number	Migrants	Natives
<b>Positive</b>		
1	fans-liverpool-biggest-league ( <i>n</i> = 291)	drivers-road-car-cycling ( <i>n</i> = 202)
2	school-schools-teachers-students ( <i>n</i> = 283)	bettermaternitycare-restrictions-maternity-partner ( <i>n</i> = 161)
3	congratulations-congrats-seanad-delighted ( <i>n</i> = 244)	ukraine-putin-russian-russia ( <i>n</i> = 146)
4	new-stats-via-follower ( <i>n</i> = 223)	read-book-piece-reading ( <i>n</i> = 126)
5	women-traveller-choice-woman ( <i>n</i> = 190)	communist-imperialist-imperialism-china ( <i>n</i> = 126)
6	irish-ireland-people-the ( <i>n</i> = 185)	beautiful-photo-lovely-pic ( <i>n</i> = 98)
7	book-hani-books-read ( <i>n</i> = 182)	energy-climate-treaty-charter ( <i>n</i> = 96)
<b>Negative</b>		
1	ireland-irish-racism-traveller ( <i>n</i> = 618)	russian-ukraine-war-putin ( <i>n</i> = 361)
2	racism-white-racist-black ( <i>n</i> = 191)	drivers-road-driver-dangerous ( <i>n</i> = 247)
3	srilanka-sri-lanka-rajapaksa ( <i>n</i> = 148)	bettermaternitycare-hse-vaccine-maternity( <i>n</i> = 135)
4	debt-climate-fuel-poverty ( <i>n</i> = 109)	climate-biodiversity-energy-planet ( <i>n</i> = 86)
5	lies-liar-threats-farright ( <i>n</i> = 93)	traveller-travellers-suicide-racism ( <i>n</i> = 80)
6	bettermaternitycare-maternity-restrictions-hse ( <i>n</i> = 82)	housing-crisis-rent-the ( <i>n</i> = 68)
7	muslim-muslims-islamophobia-attack ( <i>n</i> = 81)	racism-ireland-racist-community ( <i>n</i> = 54)
8	revolution-communism-struggle-socialism ( <i>n</i> = 80)	-
9	ukraine-war-ukrainian-russian ( <i>n</i> = 77)	-
10	asylum-seekers-deportation-centres ( <i>n</i> = 76)	-

**TABLE 6.** Sample summaries generated by ChatGPT for the tweets in the first topic in each tweets set.

<b>ireland-irish-racism-traveller (Migrants-Negative)</b>
"The text is a collection of tweets and comments about the discrimination faced by Irish Travellers and the lack of attention and action taken by the government and society to address the issue. The tweets discuss the negative impact of the lack of accommodations and the inhumane conditions faced by Irish Travellers, and the prevalence of anti-Traveller sentiment in Irish society. The tweets also express frustration at the normalization of Traveller racism and the lack of support from allies and public representatives. The text calls for more action to be taken and for people to stand in solidarity against racism towards Travellers."
<b>fans-liverpool-biggest-league (Migrants-Positive)</b>
"The text is a collection of comments and opinions on various topics related to soccer, including recent matches, players, coaches, and clubs. The fans are discussing Liverpool, United, Newcastle, Chelsea, and other teams, and expressing their opinions on different topics such as player performance, game outcomes, and manager decisions. The fans are also discussing the transfer market, including the most exciting signings, and sharing their favorite moments and goals from recent matches. Some fans express their opinions on the recent takeover"
<b>drivers-road-car-cycling (Natives-Positive)</b>
"The text is a collection of tweets discussing various topics related to transportation, urban planning, and the environment. Topics include the importance of public transport, the push to make public spaces car-free, efforts to promote safe and segregated cycling, the use of cargo bikes for businesses, and the challenges faced by people who own dogs or have children without a car. People are expressing opinions about car ownership and the barriers to cycling, and sharing their experiences of driving and parking in the city. Some are advocating for better public policy and safer walking and cycling infrastructure. They also discuss the need for conveniences like toilets and parks to be mapped and readily accessible."
<b>russian-ukraine-war-putin (Natives-Negative)</b>
"This text is a collection of comments and statements regarding the Ukraine crisis, Russia's actions, and Western responses to the situation. The comments are critical of Russia's actions, including aggression towards Ukraine and involvement in war crimes. The speakers call for a concerted European response and express disappointment in the lack of action from Western countries. They also critique the role of Western governments, including the US, and question the effectiveness of sanctions in bringing about change. Additionally, some of the speakers discuss their experiences of the war and its impact on the people of Ukraine."

against Russia (natives' topic #3 from Table 5). On the other hand, in terms of negative topics by migrants, there is a clear focus on racism with diverse themes being covered, such as a comparison of their experiences to that of Ireland's travellers' community (migrants' topic #1 from Table 5), the life of black people in Ireland (migrants' topic #2 from Table 5), Islamophobia in Ireland (migrants' topic #7 from Table 5) and finally, living conditions of asylum seekers in Ireland (migrants' topic #10 from Table 5). From a policy viewpoint, migrants and natives care about justice and fairness. However,

it is observed that natives are more concerned with a foreign sense of equality, while migrant voices focus more on domestic equality. Similarly, migrants discuss lived experiences, and natives tend to discuss more administrative issues that Ireland faces; both views offer insights into what Ireland means to both communities.

We also note migrant positive topic #7 (from Table 5) which is a discussion on the book Hani and Ishu's Guide to Fake Dating by Adiba Jaigirdar; a past pupil of a Dublin secondary school. Specifically, it is a discussion surrounding



(a) Topic fans-liverpool-biggest-league in Migrants-Positive set



(b) Topic ireland-irish-racism-traveller in Migrants-Negative set



(c) Topic drivers-road-car-cycling in Natives-Positive set



(d) Topic russian-ukraine-war-putin in Natives-Negative set

FIGURE 5. Word clouds for the first topic in each tweet set.

the LGBT themes of the book in a setting often viewed as taboo within its sociocultural setting. It is nonetheless interesting to see this discussion as more prevalent in the migrant than native communities, i.e. it not receiving the same levels of recognition given its successful delivery of a complex emotively charged aspect of adolescent life.

**E. VALUE ADDITION VIA CHATGPT SUMMARIES**

The summaries generated by ChatGPT (despite often being sampled) provide a reasonable summarisation of the topic

discourse segmented by subpopulation, i.e. migrants vs natives, and content polarity, i.e. positive vs negative. We can see very clear indications of forms of action that could be taken by Irish policymakers; here, we draw specific attention to the first and third summaries in Table 6 where ChatGPT explicitly (without explicit prompts for it to do so; although we note that more precise “prompt engineering” could further focus it to) highlights potential policy directions that could be taken on the basis of the narrative contained in the Tweet topic corpus. For completion, we also include an

irrelevant topic (the second summary) to maintain the argument for the necessity for a human-in-the-loop component of the methodology, and the fourth summary to illustrate that the summary is meaningful, as the war in Ukraine is a well-known topic that can be compared to the other summaries in terms of the type of content that can be generated. We note here too that ChatGPT was (at the time of writing) trained on data that precedes Russia's invasion of Ukraine, and as such, it should not be in a position to supplement the content it has been asked to summarise. Yet, this, in combination with the other examples (in Table 6) illustrates the potential utility of the InÉire methodology for assisting policymakers via summaries that explains the topical discourse.

#### F. LIMITATION STATEMENT

While our study provides valuable insights into the experiences of migrant communities in Ireland, it is important to acknowledge the limitations inherent in our focus on Twitter data. By solely relying on this particular platform, we recognise that our findings may not be fully representative of the diverse range of migrant experiences in the country. Several key limitations are worth highlighting:

- **Selection bias:** The use of Twitter as our primary data source may introduce a selection bias, as it only captures the perspectives of individuals who actively engage with the platform. Migrants who do not use Twitter or who are not proficient in English may be underrepresented in our sample, potentially skewing our results towards certain segments of the migrant population.
- **Digital divide:** Our study assumes that Twitter usage is evenly distributed among migrants in Ireland. However, we acknowledge that variations in digital literacy, access to technology, and socioeconomic factors may contribute to a digital divide, limiting the inclusivity of our sample.
- **Linguistic limitations:** Our analysis is based on English-language tweets, which may exclude migrants who primarily communicate in languages other than English. This linguistic limitation could restrict the representation of certain cultural groups and impact the generalizability of our findings.
- **Geographic focus:** Our study specifically focuses on the Twitter landscape in Ireland. While this provides insights into the experiences of migrants within this context, it may not capture the perspectives of migrants residing in other regions or countries.

While these limitations should be taken into account when interpreting our results, we believe that our study still provides valuable insights into the experiences and perspectives of a specific subset of migrants within the Twitter community. Future research endeavours should aim to employ diverse methodologies and data sources to capture a more comprehensive understanding of the broader migrant population in Ireland.

#### VI. CONCLUSION AND FUTURE WORK

In isolating this set of users, other more common NLP methods (e.g. topic modelling, content summarisation, sentiment analysis etc.) can be employed to reveal specific challenges or themes of interest that could be relayed to policymakers. This entails our main direction for future work (along with curating a larger user dataset): to exploit the main findings of this paper to derive thematic near real-time content that Ireland's migrant populations face in a manner that can be consumed by Irish policymakers and related agencies. As such, we view this work as a series of first steps towards isolating the migrant voice in Ireland via online social media.

This work does have some limitations that also warrant further discussion. First, we (have had to) assume that all meaningful content is expressed in English. This will likely not be the case. Yet, for the purposes of identifying users of different migration status it is; provided a user has sufficient (ca. 25 or so Tweets) English content. There is also the assumption that migrant users will comment at all on social media platforms about lived experiences. Here, we note the work of [15], [16], [17], and [18] who discuss that social platforms (like social media) are often essential for marginalised communities. We would also motivate that policymakers have very little access to such discourse, and we hope in subsequent iterations of this work to uncover something of value that can help make policies in Ireland more inclusive.

In this paper, we have argued that to understand the views of migrant populations in Ireland, we need a robust methodology to isolate the migrant voice. Twitter and other social media platforms have seen a large increase in online activism, yet it can still be difficult to differentiate between the voice of the activist and the individuals they represent. In this paper, we have illustrated that it is possible to accurately classify non-natives of Ireland with as little as 25 tweets without diving into the actual content they generate, but rather how they use language. We also argue that because this approach operates at the level of linguistic markers in the text, it should be transferable to other social media platforms, with small changes to the manner with which the text is vectorised, i.e. prepared for our voting mechanism. Future studies should aim to incorporate a wider range of data sources to capture the perspectives of migrant communities who may not be active on Twitter or proficient in English. While Twitter data provides valuable insights, it is crucial to employ complementary research methods that embrace the diversity of migrant experiences. Surveys can offer a broader understanding of the sentiments, attitudes, and lived experiences of migrants, allowing for quantitative analysis and statistical generalisation. Interviews and focus groups can provide a more nuanced exploration of individual stories, allowing researchers to delve into the intricacies of personal experiences and capture the multifaceted aspects of migration. Ethnographic approaches can provide an immersive understanding of migrant communities by observing their daily lives, interactions, and cultural practices. By combining these diverse data sources,

researchers can achieve a comprehensive understanding of the challenges, aspirations, and needs of migrant populations in Ireland, ensuring that policy-making initiatives are more inclusive and responsive to their diverse perspectives.

## REFERENCES

- [1] A. Lenhart, M. Duggan, A. Perrin, R. Stepler, H. Rainie, and K. Parker, "Teens, social media & technology overview 2015," Pew Res. Center, USA, Tech. Rep. 202.419.4372, 2015.
- [2] *Social Media Fact Sheet*, Pew Res. Center, Washington, DC, USA, 2017.
- [3] K. Schiffer and E. Schatz, "Marginalisation, social inclusion and health. Experiences based on the work of correlation-European network social inclusion & health," Found. Regenboog AMOC, Amsterdam, The Netherlands, Tech. Rep., 2008.
- [4] F. R. Siddiqui, "Annotated bibliography on participatory consultations to help aid the inclusion of marginalized perspectives in setting policy agendas," *Int. J. Equity Health*, vol. 13, no. 1, pp. 1–16, Dec. 2014.
- [5] B. Fanning, "Multiculturalism in Ireland," in *Racism and Social Change in the Republic of Ireland*. Manchester, U.K.: Manchester University Press, 2018, pp. 178–198.
- [6] The Department of Justice. (2022). *Equality, Diversity and Inclusion Strategy*. [Online]. Available: [https://www.justice.ie/en/JELR/Pages/DoJ\\_EDI\\_Strategy](https://www.justice.ie/en/JELR/Pages/DoJ_EDI_Strategy)
- [7] L. Michael. (2020). *Reports of Racism in Ireland*. [Online]. Available: [https://inar.ie/wp-content/uploads/2020/03/2019\\_iReport\\_Final.pdf](https://inar.ie/wp-content/uploads/2020/03/2019_iReport_Final.pdf)
- [8] M. Gońda, M. Lesińska, and M. Pachocka, "Relations between immigration and integration policies in postwar Europe," *Relations Between Immigration and Integration Policies in Europe*. London, U.K.: Routledge, 2020, p. 24.
- [9] R. O. Lutkenhaus, J. Jansz, and M. P. A. Bouman, "Mapping the Dutch vaccination debate on Twitter: Identifying communities, narratives, and interactions," *Vaccine: X*, vol. 1, Apr. 2019, Art. no. 100019.
- [10] J. Kim, A. Sirbu, F. Giannotti, and L. Gabrielli, "Digital footprints of international migration on Twitter," in *Proc. Int. Symp. Intell. Data Anal.* Cham, Switzerland: Springer, 2020, pp. 274–286.
- [11] M. Mazzoli, B. Diechtiareff, A. Tugores, W. Wives, N. Adler, P. Colet, and J. J. Ramasco, "Migrant mobility flows characterized with digital data," *PLoS ONE*, vol. 15, no. 3, 2020, Art. no. e0230264.
- [12] A. Younus, M. A. Qureshi, M. Jeon, A. Kazemi, and S. Caton, "XAI analysis of online activism to capture integration in Irish society through Twitter," in *Proc. Int. Conf. Social Inform.* Springer, 2022, pp. 233–244.
- [13] M. Hall and S. Caton, "Am i who i say i am? Unobtrusive self-representation and personality recognition on Facebook," *PLoS ONE*, vol. 12, no. 9, Sep. 2017, Art. no. e0184417.
- [14] C. Andrews, "Social media recruitment: The opportunities and challenges social media provides in the realm of subject recruitment," *Appl. Clin. Trials*, vol. 21, no. 11, p. 32+, Nov. 2012.
- [15] A. L. Gonzales, "Disadvantaged minorities' use of the internet to expand their social networks," *Commun. Res.*, vol. 44, no. 4, pp. 467–486, Jun. 2017.
- [16] M. A. Devito, A. M. Walker, J. Birnholtz, K. Ringland, K. Macapagal, A. Kraus, S. Munson, C. Liang, and H. Saksono, "Social technologies for digital wellbeing among marginalized communities," in *Proc. Conf. Companion Publication Comput. Supported Cooperat. Work Social Comput.*, Nov. 2019, pp. 449–454.
- [17] J. Fox and R. Ralston, "Queer identity online: Informal learning and teaching experiences of LGBTQ individuals on social media," *Comput. Hum. Behav.*, vol. 65, pp. 635–642, Dec. 2016.
- [18] D. Nemer, "Online favela: The use of social media by the marginalized in Brazil," *Inf. Technol. Develop.*, vol. 22, no. 3, pp. 364–379, Jul. 2016.
- [19] K. Bender, S. Begun, A. DePrince, B. Haffeejee, and S. Kaufmann, "Utilizing technology for longitudinal communication with homeless youth," *Social Work Health Care*, vol. 53, no. 9, pp. 865–882, Oct. 2014.
- [20] S. Caton, M. Hall, and C. Weinhardt, "How do politicians use Facebook? An applied social observatory," *Big Data Soc.*, vol. 2, no. 2, 2015, Art. no. 2053951715612822.
- [21] P. Burnap, M. L. Williams, L. Sloan, O. Rana, W. Housley, A. Edwards, V. Knight, R. Procter, and A. Voss, "Tweeting the terror: Modelling the social media reaction to the Woolwich terrorist attack," *Social Netw. Anal. Mining*, vol. 4, no. 1, pp. 1–14, Dec. 2014.
- [22] D. Kilroy, S. Caton, and G. Healy, "Finding short lived events on social media," in *Proc. 28th AIAI Irish Conf. Artif. Intell. Cogn. Sci. (AICS)*, 2020, pp. 1–12.
- [23] L. M. Aiello, G. Petkos, C. Martin, D. Corney, S. Papadopoulos, R. Skraba, A. Göker, I. Kompatsiaris, and A. Jaimes, "Sensing trending topics in Twitter," *IEEE Trans. Multimedia*, vol. 15, no. 6, pp. 1268–1282, Oct. 2013.
- [24] T. de Melo and C. M. S. Figueiredo, "A first public dataset from Brazilian Twitter and news on COVID-19 in Portuguese," *Data Brief*, vol. 32, Oct. 2020, Art. no. 106179.
- [25] G. A. Irgens, "Using knowledgeable agents of the digital and data feminism to uncover social identities in the #blackgirlmagic Twitter community," *Learn. Media Technol.*, vol. 47, no. 1, pp. 79–94, Jan. 2022.
- [26] F. Poletto, V. Basile, M. Sanguinetti, C. Bosco, and V. Patti, "Resources and benchmark corpora for hate speech detection: A systematic review," *Lang. Resour. Eval.*, vol. 55, no. 2, pp. 477–523, Jun. 2021.
- [27] U. Naseem, I. Razzak, and P. W. Eklund, "A survey of pre-processing techniques to improve short-text quality: A case study on hate speech detection on Twitter," *Multimedia Tools Appl.*, vol. 80, nos. 28–29, pp. 35239–35266, Nov. 2021.
- [28] E. Calvete, I. Orue, and M. Gámez-Guadix, "Cyberbullying victimization and depression in adolescents: The mediating role of body image and cognitive schemas in a one-year prospective study," *Eur. J. Criminal Policy Res.*, vol. 22, no. 2, pp. 271–284, Jun. 2016.
- [29] P. Burnap and M. L. Williams, "Cyber hate speech on Twitter: An application of machine classification and statistical modeling for policy and decision making," *Policy Internet*, vol. 7, no. 2, pp. 223–242, Jun. 2015.
- [30] A. Matamoros-Fernández and J. Farkas, "Racism, hate speech, and social media: A systematic review and critique," *Telev. New Media*, vol. 22, no. 2, pp. 205–224, Feb. 2021.
- [31] S. Caton, "Using social media and text analytics to investigate marginal phenomena," Eurofound, Dublin, Ireland, Working Paper WPEF21052, 2022.
- [32] D. U. Patton, J. Lane, P. Leonard, J. Macbeth, and J. R. Smith Lee, "Gang violence on the digital street: Case study of a south side Chicago gang member's Twitter communication," *New Media Soc.*, vol. 19, no. 7, pp. 1000–1018, Jul. 2017.
- [33] M. Nartey, "Centering marginalized voices: A discourse analytic study of the black lives matter movement on Twitter," *Crit. Discourse Stud.*, vol. 19, no. 5, pp. 523–538, 2021.
- [34] K. Coe and R. A. Griffin, "Marginalized identity invocation online: The case of president Donald Trump on Twitter," *Social Media+ Soc.*, vol. 6, no. 1, 2020, Art. no. 2056305120913979.
- [35] D. Murthy, A. Gross, and A. Pensavalle, "Urban social media demographics: An exploration of Twitter use in major American cities," *J. Comput.-Mediated Commun.*, vol. 21, no. 1, pp. 33–49, Jan. 2016.
- [36] R. T. Ionescu, M. Popescu, and A. Cahill, "Can characters reveal your native language? A language-independent approach to native language identification," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2014, pp. 1363–1373.
- [37] B. Batrinca and P. C. Treleaven, "Social media analytics: A survey of techniques, tools and platforms," *AI Soc.*, vol. 30, no. 1, pp. 89–116, Feb. 2015.
- [38] M. Hall, A. Mazarakis, M. Chorley, and S. Caton, "Editorial of the special issue on following user pathways: Key contributions and future directions in cross-platform social media research," *Int. J. Hum.-Comput. Interact.*, vol. 34, no. 10, pp. 895–912, Oct. 2018.
- [39] D. Rogers, A. Preece, M. Innes, and I. Spasic, "Real-time text classification of user-generated content on social media: Systematic review," *IEEE Trans. Computat. Social Syst.*, vol. 9, no. 4, pp. 1154–1166, Aug. 2022.
- [40] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems?" *J. Mach. Learn. Res.*, vol. 15, pp. 3133–3181, Jan. 2014.
- [41] S. A. Salloum, M. Al-Emran, A. A. Monem, and K. Shaalan, "A survey of text mining in social media: Facebook and Twitter perspectives," *Adv. Sci., Technol. Eng. Syst. J.*, vol. 2, no. 1, pp. 127–133, Jan. 2017.
- [42] J. Garland, K. Ghazi-Zahedi, J.-G. Young, L. Hébert-Dufresne, and M. Galesic, "Countering hate on social media: Large scale classification of hate and counter speech," in *Proc. 4th Workshop Online Abuse Harms*, 2020, pp. 102–112.
- [43] P. Burnap, G. Colombo, R. Amery, A. Hodorog, and J. Scourfield, "Multi-class machine classification of suicide-related communication on Twitter," *Online Social Netw. Media*, vol. 2, pp. 32–44, Aug. 2017.
- [44] S. Ahmad, M. Z. Asghar, F. M. Alotaibi, and I. Awan, "Detection and classification of social media-based extremist affiliations using sentiment analysis techniques," *Human-Centric Comput. Inf. Sci.*, vol. 9, no. 1, pp. 1–23, Dec. 2019.
- [45] S. G. Burdisso, M. Errecalde, and M. Montes-Y-Gómez, "A text classification framework for simple and effective early depression detection over social media streams," *Exp. Syst. Appl.*, vol. 133, pp. 182–197, Nov. 2019.

- [46] M. Chau, T. M. H. Li, P. W. C. Wong, J. J. Xu, P. S. F. Yip, and H. Chen, "Finding people with emotional distress in online social media: A design combining machine learning and rule-based classification," *MIS Quart.*, vol. 44, no. 2, pp. 933–955, Jun. 2020.
- [47] F. Huang, "Improved Arabic dialect classification with social media data," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2015, pp. 2118–2126.
- [48] D. Blanchard, J. Tetreault, D. Higgins, A. Cahill, and M. Chodorow, "TOEFL11: A corpus of non-native English," *ETS Res. Rep. Ser.*, vol. 2013, no. 2, p. 15, Dec. 2013.
- [49] E. Rabinovich, Y. Tsvetkov, and S. Wintner, "Native language cognate effects on second language lexical choice," *Trans. Assoc. Comput. Linguistics*, vol. 6, pp. 329–342, Dec. 2018.
- [50] S. Volkova, S. Ranshous, and L. Phillips, "Predicting foreign language usage from English-only social media posts," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Human Lang. Technol.*, 2018, pp. 608–614.
- [51] G. Goldin, E. Rabinovich, and S. Wintner, "Native language identification with user generated content," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2019, pp. 3591–3601.
- [52] S. Steinbakken and B. Gambäck, "Native-language identification with attention," in *Proc. 17th Int. Conf. Natural Lang. Process. (ICON)*, 2020, pp. 261–271.
- [53] R. Irfan, C. K. King, D. Grages, S. Ewen, S. U. Khan, S. A. Madani, J. Kolodziej, L. Wang, D. Chen, A. Rayes, N. Tziritas, C.-Z. Xu, A. Y. Zomaya, A. S. Alzahrani, and H. Li, "A survey on text mining in social networks," *Knowl. Eng. Rev.*, vol. 30, no. 2, pp. 157–170, 2015.
- [54] J. W. Pennebaker, M. E. Francis, and R. J. Booth, "Linguistic inquiry and word count: LIWC 2001," *Mahway: Lawrence Erlbaum Associates*, vol. 71, no. 2001, p. 2001, 2001.
- [55] J. W. Pennebaker, R. L. Boyd, K. Jordan, and K. Blackburn, "The development and psychometric properties of LIWC2015," Univ. Texas Austin, Austin, TX, USA, 2015. [Online]. Available: [https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015\\_LanguageManual.pdf?sequence=3&isAllowed=y](https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015_LanguageManual.pdf?sequence=3&isAllowed=y)
- [56] A. Balahur and J. M. Hermida, "Extending the emotinet knowledge base to improve the automatic detection of implicitly expressed emotions from text," in *Proc. LREC*, 2012, pp. 1207–1214.
- [57] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," in *Proc. 49th Annu. Meeting Assoc. Comput. Linguistics, Human Lang. Technol.*, 2011, pp. 309–319.
- [58] M. L. Newman, J. W. Pennebaker, D. S. Berry, and J. M. Richards, "Lying words: Predicting deception from linguistic styles," *Personality Social Psychol. Bull.*, vol. 29, no. 5, pp. 665–675, May 2003.
- [59] Y. R. Tausczik and J. W. Pennebaker, "The psychological meaning of words: LIWC and computerized text analysis methods," *J. Lang. Social Psychol.*, vol. 29, no. 1, pp. 24–54, Mar. 2010.
- [60] S. González-Bailón, N. Wang, A. Rivero, J. Borge-Holthoefer, and Y. Moreno, "Assessing the bias in samples of large online networks," *Social Netw.*, vol. 38, pp. 16–27, Jul. 2014.
- [61] B. A. Nosek, M. R. Banaji, and A. G. Greenwald, "E-research: Ethics, security, design, and control in psychological research on the internet," *J. Social Issues*, vol. 58, no. 1, pp. 161–176, Jan. 2002.
- [62] A. Zwitter, "Big data ethics," *Big Data Soc.*, vol. 1, no. 2, 2014, Art. no. 2053951714559253.
- [63] D. Boyd and K. Crawford, "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon," *Inf., Commun. Soc.*, vol. 15, no. 5, pp. 662–679, Jun. 2012.
- [64] D. Gaffney and J. N. Matias, "Caveat emptor, computational social science: Large-scale missing data in a widely-published Reddit corpus," *PLoS ONE*, vol. 13, no. 7, Jul. 2018, Art. no. e0200162.
- [65] M. Rost, L. Barkhuus, H. Cramer, and B. Brown, "Representation and communication: Challenges in interpreting large social media datasets," in *Proc. Conf. Comput. Supported Cooperat. Work*, Feb. 2013, pp. 357–362.
- [66] L. Townsend and C. Wallace, "Social media research: A guide to ethics," Univ. Aberdeen, Aberdeen, U.K., Tech. Rep., 2016.
- [67] A. Markham, "Ethical decision-making and internet research: Version 2.0," Assoc. Internet Researchers, Philadelphia, PA, USA, Tech. Rep., 2012.
- [68] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, A. Flammini, and F. Menczer, "Detecting and tracking political abuse in social media," in *Proc. Int. AAAI Conf. Web Social Media*, vol. 5, no. 1, 2011, pp. 297–304.
- [69] P. T. Metaxas and E. Mustafaraj, "Social media and the elections," *Science*, vol. 338, no. 6106, pp. 472–473, 2012.
- [70] A. Gupta, H. Lamba, and P. Kumaraguru, "\$1.00 per RT# boston-marathon# prayforboston: Analyzing fake content on Twitter," in *Proc. APWG eCrime Researchers Summit*, 2013, pp. 1–12.
- [71] C. Cassa, R. Chunara, K. Mandl, and J. S. Brownstein, "Twitter as a sentinel in emergency situations: Lessons from the Boston Marathon explosions," *PLoS Currents*, vol. 5, May 2013.
- [72] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," *Commun. ACM*, vol. 59, no. 7, pp. 96–104, Jun. 2016.
- [73] T. Hagerup, K. Mehlhorn, and J. I. Munro, "Maintaining discrete probability distributions optimally," in *Proc. 20th Int. Colloq. Automata, Lang. Program.*, in Lecture Notes in Computer Science, vol. 700, Berlin, Germany: Springer-Verlag, 1993, pp. 253–264.
- [74] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *Proc. Int. AAAI Conf. Web Social Media*, 2014, vol. 8, no. 1, pp. 216–225.
- [75] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Jan. 2003.
- [76] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," 2022, *arXiv:2203.05794*.



**AREFEH KAZEMI** received the B.S. and M.S. degrees in software engineering and the Ph.D. degree in artificial intelligence from the University of Isfahan, Isfahan, Iran, in 2008, 2010, and 2017, respectively. She is currently a level 2 Postdoctoral Researcher with the Computer Science Department, University College Dublin. From 2019 to 2022, she was an Assistant Professor with the Computational Linguistics Group, University of Isfahan. Her research interests include machine learning, natural language processing, data mining, and applied artificial intelligence.



**ARJUMAND YOUNUS** received the B.S. degree in computer science from the University of Karachi, Pakistan, the M.S. degree in computer science from the Korea Advanced Institute of Science and Technology, South Korea, and the joint Ph.D. degree in computer science from the National University of Ireland Galway, Ireland, and University of Milano-Bicocca, Italy. She is currently an Assistant Professor in computational social science with the School of Sociology, University College Dublin. Formerly, she was a Research Scientist with Afiniti AI and a part-time Lecturer with Technological University Dublin. She was a recipient of Google Women Techmakers Scholarship for Europe, Middle East and Africa region. She is also serving in the capacity of the Co-Director for Women in Research Ireland, which is a volunteer-run registered charity working for better representation of women and underrepresented groups in academia. Her research interests include machine learning, natural language processing, and data science for social good. She is passionate about the value of artificial intelligence technology to make society better, and at the moment, is involved as an academic partner in various AI for social good projects.





mobility, technological learning, exploratory and exploitative innovation, and cultural distance.

**MINGYEONG JEON** received the Ph.D. degree in international business from the Leeds University Business School, U.K., in 2021. She is currently a Postdoctoral Researcher with Technological University Dublin, Ireland. From 2015 to 2017, she was a Research Assistant with the Korea Labor Institute and an Associate Researcher with the Korea Information Society Development Institute, South Korea. Her research interests include cross-border knowledge spillover through human



include the applications of machine learning across the domains of social media, quantum computing, and parallel and distributed computing.

**SIMON CATON** received the B.Sc. (Hons.) and Ph.D. degrees in computer science from Cardiff University, U.K., in 2005 and 2010, respectively. He was a Postdoctoral Researcher with the Karlsruhe Institute of Technology, Germany, from 2010 to 2014, and a Lecturer in data analytics with the National College of Ireland, from 2014 to 2019. He is currently an Assistant Professor with the School of Computer Science, University College Dublin. His research interests

...



interests include natural language processing, machine learning, disinformation space, and explainable artificial intelligence. He has contributed to various projects funded by Science Foundation Ireland, Irish Research Council, Enterprise Ireland, and the EU as a Principal Investigator and a Technical Lead and licensed an outcome to a leading media organization in Ireland. He is passionate about applied research of interdisciplinary nature.

**M. ATIF QURESHI** received the B.S. degree in computer science from the University of Karachi, Pakistan, the M.S. degree in computer science from the Korea Advanced Institute of Science and Technology, South Korea, and the joint Ph.D. degree in computer science from the National University of Ireland Galway, Ireland, and the University of Milano-Bicocca, Italy. He is currently a Lecturer/an Assistant Professor with Technological University Dublin, Ireland. His research