



**HAL**  
open science

## Public Perception towards fifth generation of cellular networks (5G) on social media

Kia Dashtipour, William Taylor, Shuja Ansari, Mandar Gogate, Adnan Zahid, Yusuf Sambo, Amir Hussain, Qammer H Abbasi, Muhammad Ali Imran

### ► To cite this version:

Kia Dashtipour, William Taylor, Shuja Ansari, Mandar Gogate, Adnan Zahid, et al.. Public Perception towards fifth generation of cellular networks (5G) on social media. *Frontières*, 2021. hal-03217618

**HAL Id: hal-03217618**

**<https://hal.science/hal-03217618>**

Submitted on 4 May 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Public Perception towards fifth generation of cellular networks (5G) on social media

Kia Dashtipour<sup>1,\*</sup>, William Taylor<sup>1</sup>, Shuja Ansari<sup>1</sup>, Mandar Gogate<sup>2</sup>, Adnan Zahid<sup>1</sup>, Yusuf Sambo<sup>1</sup>, Amir Hussain<sup>2</sup>, Qammer H. Abbasi<sup>1</sup> and Muhammad Ali Imran<sup>1</sup>

<sup>1</sup> James Watt School of Engineering, University of Glasgow, Glasgow, G12 8QQ, U.K.

<sup>2</sup> School of Computing, Edinburgh Napier University, Edinburgh EH10 5DT, UK

Correspondence\*:

Corresponding Author

kia.dashtipour@glasgow.ac.uk

## 2 ABSTRACT

3 With advancement of social media network, there are lots of unlabelled reviews available online,  
4 therefore its necessarily to develop an automatic tools to classify these types of reviews. For  
5 utilising these reviews for user perception, there is a need for automated tools that can process  
6 online user data for optimising user perception. In this paper, a sentiment analysis framework has  
7 been proposed to identify people's perception towards mobile networks. The proposed framework  
8 consists of three basic steps: preprocessing, feature selection and applying different machine  
9 learning algorithms. The performance of the framework has taken into account different feature  
10 combinations. The simulation results show that the best performance is by integrating unigram,  
11 bigram and trigram features.

12 **Keywords:** Sentiment Analysis, 5G, Mobile Network Quality, Machine Learning, Opinion Mining

## 1 INTRODUCTION

13 The 5th generation mobile network is new global wireless standard after 1G, 2G, 3G and 4G network. The  
14 5G is new types of network which is designed and developed to connect virtually everyone and everything  
15 together and consists of different machines, objectives and devices. In addition, the fifth generation of  
16 cellular networks is required to be more efficient and economical in terms of Key Performance Indicators  
17 (KPI). KPIs are of interest to stakeholders and different applications. These KPIs from an operator  
18 perspective, consist of capacity, reliability and quality of service. From the user perspective, the KPIs  
19 include uninterrupted connection, infinite capacity and zero latency. However, there no technology that  
20 can offer infinite capacity or zero latency. In the past few years, there has been lots of research carried out  
21 in next generation mobile network (5G), which consists of different opportunities and challenges. There  
22 are different challenges in fifth generation of cellular networks (5G) that have been discussed in different  
23 literature. The most significant of these are ultra-dense networks and millimetre waves, however, there are  
24 other technologies which are significant for the next generation of network such as two-layer architecture  
25 and cognitive radio based architectures which have great performance in the next generation of the network.  
26 In fifth generation of cellular networks (5G), network data analytic and machine learning systems can  
27 perform a key role. The technique to understand people's behaviour towards the 5G network is vital and it

28 helps to improve the performance of the network communication Sharma et al. (2020) OPINCARIU et al.  
29 (2019).

30 With the advent of social media and e-commerce, websites allow users to share opinions and feedback  
31 about different products and services. Customers can make important decisions by reading other people's  
32 experiences. In addition, the customer feedback can be classified to make improvements on the service or  
33 product. For example, if a person wants to buy a mobile phone and the reviews provide negative information  
34 related to the battery, operation speed or camera, this can influence the consumers decision. In addition, this  
35 can assist in providing better mobile quality by taking into account the complaints made by past customers  
36 and making informed improvements to products. As another example, if someone wants to book a hotel,  
37 the buyer can look over the online reviews to understand previous customer experience such as cleanliness  
38 and services for the hotel. However, there are billions of bytes of data generated per day consisting of user  
39 feedback which cannot be manually labelled and analysed for individual organisations and companies  
40 Yadav and Vishwakarma (2020). Sentiment analysis is the process of automatically understanding and  
41 classifying the data into positive and negative information from the source material such as reviews and  
42 comments. The main task of sentiment analysis is to assign polarity into sentences (positive or negative).  
43 However, the online review is a mixture of positive and negative comments about different aspects of the  
44 products or services instead of expressing positive or negative opinions. For example, "the 5G mobile  
45 network is extremely fast, however I do not feel secure while I am using the 5G network". The sentence  
46 expresses positive sentiment towards the speed of the mobile and negative sentiment towards the security  
47 of the mobile network Kumar and Jaiswal (2020), Dashtipour et al. (2020).

48 Most of the current sentiment analysis approaches are focused on analysing products and movie reviews  
49 and there is less work been carried out in different fields such as people's perception towards fifth  
50 generation of cellular networks (5G). However, most of the current approaches consider a small corpus,  
51 which makes the task difficult for machine learning approaches to identify the overall polarity of the  
52 sentence. Furthermore, the current sentences consist of lots of sarcastic and ironic words which is difficult  
53 to determine the overall polarity for these types of sentences. For example, "Tell me something that I do not  
54 know". Most of the current approaches for sentiment analysis fail to understand the real noisy text consists  
55 of sarcasm, idioms, informal words and sentences with spelling mistakes. In addition, there are scarce  
56 availability of tools and resources. lexicon and labelled corpus are some of the tools which are available  
57 for sentiment analysis. This limited range of available tools is the main bottleneck in the design of the  
58 sentiment analysis approaches Kaity and Balakrishnan (2019). One of the main issue for sentiment analysis  
59 approaches is lack of labelled dataset. However, it is worth to mention that, there are lots of unlabelled  
60 dataset available online, but it is time consuming for users to manually label dataset.

61 In order to address the aforementioned issues and limitations, framework that exploits n-gram features has  
62 been proposed to identify the polarity of sentence. This proposed approach demonstrated that the overall  
63 performance and effectiveness of polarity detection in real noisy data. The ngram features are based on the  
64 linguistic text rules that allow to extract text features from sentences. As a result, the ngram features take  
65 into account the relation between keywords and the word order and individual word polarity to determine  
66 the underlying polarity of the sentence. We perform an extensive and comprehensive set of experiments  
67 using novel corpus and compare the performance of the approach with different selected features. The  
68 Support Vector Machine(SVM), logistic regression, Naive Bayes and Multi Layer Perceptron (MLP) used  
69 to evaluate the performance of the approach. The comparative simulation results show that the proposed  
70 approach achieved better performance as compared to the state-of-art approach.

71 The rest of the paper is organised as follows: In Section 2 related work is presented, Section 3 presented  
72 the proposed framework, Section 4 presented the experimental results, Section 5 presents discussion and  
73 finally Section 6 concludes the work and presents the future work.

## **2 RELATED WORK**

74 The extensive research in the current literature shows that the machine learning has been used in different  
75 fields such as sentiment analysis Hussain et al. (2021); Dashtipour et al. (2021, 2016b,a, 2017c,b,a); Jiang  
76 et al. (2020); Dashtipour et al. (2018, 2020); Guellil et al. (2021); Dashtipour et al. (2019), cyber-security  
77 Adeel et al. (2019a); Jiang et al. (2019); Shiva et al. (2017); Gogate et al. (2019b), hand-written recognition  
78 Ahmed et al. (2021), speech enhancement Gogate et al. (2020a,b, 2019a, 2017a, 2018); Ozturk et al. (2019);  
79 Adeel et al. (2020, 2019b); Gogate et al. (2017b), posture detection Taylor et al. (2020); Liaqat et al. (2021,  
80 2020) and etc. However, there is not any research carried out to detect sentiment polarity for tweets related  
81 to 5G.

82 In the literature, extensive research has been carried out to implement different sentiment analysis  
83 approaches. The microblogging websites are the biggest platform that allows users to share their thoughts  
84 and opinions in the public domain. Twitter is the most well-known microblogging website that allows  
85 people to express their feelings and emotions in the form of "tweets" with a character range of 280. There  
86 are over 250 million tweets expressing the feelings and emotions of people with different opinions and  
87 situations Duong et al. (2019) Mamgain et al. (2016).

88 The twitter users vary from politicians to everyday people that provide different types of reviews from  
89 different points of views. This is the the main reason a dataset collected from twitter is used in this paper.  
90 There is lots of research carried out in this domain to determine people's perception towards different  
91 products such as Sony mobile. For example. Sharma et al. (2016) focused to find the most well known  
92 smartphones in India, the tweets are collected and then machine learning is applied to determine a brand  
93 reputation score. This will help customers to find the most branded smartphones in India. Somula et al.  
94 (2020) proposed an approach to perform a sentiment analysis to determine the winner of the US election  
95 in 2016, the tweets are collected with respect to Donald Trump and Hillary Clinton, the tweets analysing  
96 results reveals that Donald Trump received more positive scores as compared to Hillary Clinton. Mehta  
97 et al. (2020) proposed a sentiment analysis approach to identify the best Indian airlines through twitter, the  
98 analysis of results reveals that the customers are happier with airIndia services compared to spicejet.

99 Kumari and Haider (2020) proposed an approach using twitter API to collect corpus, after pre-processing  
100 and use of natural language processing, the hybrid classifier utilised machine learning and long short-term  
101 memory to improve the performance of the approach. The experimental results reveal that the proposed  
102 model achieved better performance as compared with state-of-the-art approaches. Usama et al. (2019)  
103 introduced a novel model multi-level feature extraction and feature combination by using convolutional  
104 neural network (CNN) and recurrent neural network (RNN) to identify the sentiment in movie reviews.  
105 The CNN and RNN received sentiment text as input and learn different features to network architecture.  
106 The word embedding fed into CNN and learn multilevel contextual features from every layer of CNN and  
107 perform multilevel features fusion. Finally, the multilevel and multitype features are combined and softmax  
108 classifier is used to identify final polarity of sentences. Most of the aforementioned studies implement a  
109 lexicon to determine the polarity of the text. However, the use of lexicon fails to identify the polarity of  
110 word order in the sentence. In addition, most of the current studies use rules to detect negation in a sentence  
111 which cannot be directly applied to English language. However, we need a framework to identify the  
112 polarity of the sentence without developing a lexicon which is time consuming. Therefore, we implemented

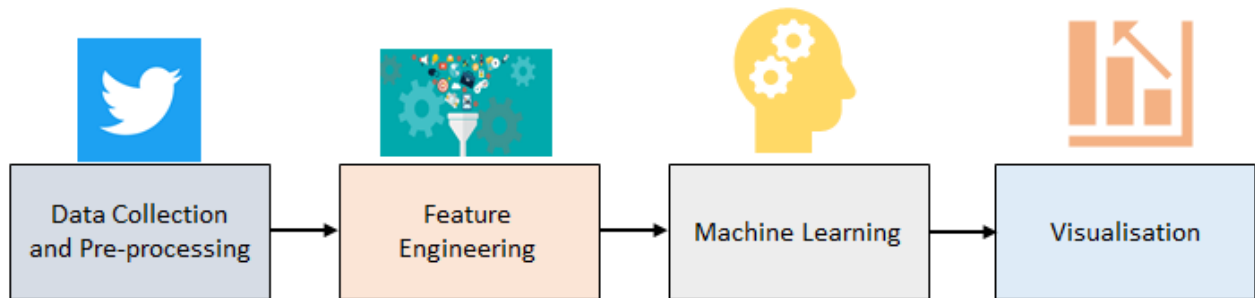
**Table 1.** Next Generation Mobile Network (5G)

Keywords
5G
Next generation mobile network
fifth generation of technology
5G devices

113 a framework for twitter sentiment analysis that integrates feature engineering and machine learning to  
 114 improve the performance and robustness of polarity detection in real noisy data.

### 3 METHODOLOGY

115 This section describes our proposed novel context-aware framework for fifth generation of cellular networks  
 116 (5G) sentiment analysis. The proposed framework exploits the polarity of sentence more accurately as  
 117 compared to traditional word occurrence frequency-based approaches.



**Figure 1.** Proposed Framework for Sentiment Analysis for people perception towards fifth generation of cellular networks (5G)

118 **Data Collection:** In order to collect data, we used Twitter API to collect data related to next generation  
 119 mobile network (5G) in the UK. The data is collected from January 2018 till August 2020. The Table 1  
 120 shows the related keywords for fifth generation of cellular networks (5G).

121 **Pre-processing:** The tweets corpus has been collected using twitter API and it has been labelled into  
 122 positive and negative using SentiWordNet. The corpus is divided into a training set (60%), test set (30%)  
 123 and validation set (10%) to apply machine learning algorithms including SVM, Naive Bayes and MLP. The  
 124 corpus is tokenized and normalised. The tokenisation technique is used to break sentences into words. For  
 125 example, “I really like mobile” will be converted into words such as “I”, “really”, “like” and “mobile”.  
 126 Afterwards, the normalisation technique is used to normalise the tweets. For example, “The speed for 5G  
 127 mobile is gr8” will be converted into “The speed for 5G mobile is great”.

128 **Ngram:** N-gram features are widely used in different approaches of sentiment analysis. When there is  
 129 one term is taken as feature is called unigram, for two term is called bigram and three terms is called  
 130 trigram. In our proposed approach we have been using unigram, bigram and trigram and we used the  
 131 combination of n-gram features.

132 **SentiWordNet:** In order to assign sentiment polarity into sentence we used SentiWordNet widely available  
 133 online lexicon to assign sentiment polarity (-1, 0, 1) into sentences.

134 **Machine learning classifiers:** In order to evaluate the performance of the approach, the machine learning  
 135 classifier is used to evaluate the performance of the approach. The scikit-learn python package is used to  
 136 developed multi-layer perceptron, logistic regression, linear SVM, RBF SVM and naive bayes is used to  
 137 train the model. The MLP consist of one hidden layer which is able to apply to supervised problem, the  
 138 MLP are set of input and put and it learns to model the correlation between input and output. However, the  
 139 main issue with MLP it consists of many parameters which is fully connected and each nodes is connected  
 140 to other node which can result in redundancy and inefficiency. The main advantage of Naïve Bayes is  
 141 performs fast and save lots of time however, the prediction can be wrong. In addition, the main advantage  
 142 of logistic regression is easy to implement and train however, the number of observation is less than number  
 143 of features. Finally, the main advantages of SVM is work well with unstructured data such as text, as the  
 144 tweets are unstructured, therefore, the SVM performs well of tweets data. However, the main issue with  
 145 SVM, the choosing correct kernel is not easy.

#### 4 EXPERIMENTAL RESULTS

146 In order to calculate the sentiment polarity of the sentence, the SentiWordNet has been used to calculate  
 147 overall polarity of the tweets. The tweets extracted using different keywords such as 5G, next generation  
 148 mobile network, fifth generation of technology and 5g devices. The tweets are collected. After tweets  
 149 collection, the punctation and stop words are removed and sentence are normalised, and then its been  
 150 converted into BOW and finally machine learning classifiers trained. In the pre-processing stage, the  
 151 sentence are normalised, for example, the word "fishing" are converted into "fish".

152 In order to evaluate the performance of the proposed approach, the tweets are converted into bag of words  
 153 (BOW). These BOWs are sent to machine learning algorithms including Linear and RBF SVM, naive  
 154 bayes, logistic regression and MLP to evaluate the performance of the approach. The intial experimental  
 155 results demonstrate that the combination of unigram and bigram achieved accuracy of 86.71%.

156 **Dataset:** In order to evaluate the performance of the approach the tweets have been collected from twitter,  
 157 the 5G hashtag has been used to collect tweets, more than fifty thousand tweets have been collected and the  
 158 positive and negative polarity has been assigned into tweets using SentiWordNet lexicon. The neural tweets  
 159 have been eliminated. The experimental results show that the combination of unigram and bigram achieved  
 160 better performance as compared to other approaches. In order to evaluate the performance of the proposed  
 161 approach, there are different evaluation metrics including accuracy, precision, recall and f-measure are  
 162 used:

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F\_measure = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (3)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

**Table 2.** Paramets of ML algorithms

Algorithm	Parameter	Time
SVM	RBF Kernal	4 mins and 21 s
Naive Bayes	sample weight = none	2 mins and 12 s
MLP	activation = relu	3 mins and 31 s
Logistic regression	Penalty = l2	3 mins 42 s

**Table 3.** Results of N-gram Features

Feature	Classifier	Accuracy	Precision	Recall	F-score
Uni	MLP	85.92	0.86	0.86	0.86
Uni	LR	86.14	0.86	0.86	0.86
Uni	Linear SVM	84.79	0.84	0.84	0.84
Uni	RBF SVM	62.61	0.72	0.63	0.54
Uni	NB	86.48	0.87	0.86	0.87
Bi	MLP	79.16	0.8	0.79	0.78
Bi	LR	78.37	0.79	0.78	0.78
Bi	Linear SVM	73.98	0.78	0.74	0.72
Bi	RBF SVM	62.61	0.72	0.63	0.54
Bi	NB	78.82	0.8	0.79	0.78
Tri	MLP	73.42	0.77	0.73	0.71
Tri	LR	72.52	0.77	0.73	0.7
Tri	Linear SVM	69.48	0.77	0.69	0.65
Tri	RBF SVM	62.61	0.72	0.63	0.54
Tri	NB	71.39	0.77	0.71	0.68

166 where TP denotes true positive, TN presents true negative, FP is false positive, and FN represents false  
 167 negative respectively. In addition, Table 2 shows the parameters that are used to trained the machine  
 168 learning methods. The scikit-learn package is used to train the machine learning classifiers. In addition, the  
 169 training time for each models has been presented in Table 2.

170 Table 3 shows the results of different N-gram features. The unigram (Uni), bigram (Bi) and trigram (Tri)  
 171 is extracted from the sentence, as the comparative experimental result shows the unigram (Uni) achieved  
 172 better performance as compared to other features.

173 Table 4 shows the comparison of different N-gram features. The empirical results shows that the  
 174 combination unigram (Uni) and bigram (Bi) is achieved better performance as compared to other features.

175 Figure 2 displays the positive sentiment towards fifth generation of cellular networks (5G) technology, the  
 176 speed, security and performance towards 5G. For example, the user has a positive opinion towards speed,  
 177 security and performance. In order to find the most positive keywords towards fifth generation of cellular  
 178 networks (5G) the word frequency of the words positive sentences is calculated. The word frequency shows  
 179 that the most discussed keywords are speed, security and performance. As shown in Figure 2 (a) Scotland  
 180 and England have the most positive discussion towards speed. For example, 45% of tweets in Scotland and  
 181 30% in England has positive opinion about the speed of 5G. In the other hand, Figure 2 (b) demonstrate that  
 182 England and Scotland have most positive discussion toward security trend. 41% of tweets in England and  
 183 36% in Scotland has positive opinion about the security of 5G. Figure 2 (c) shows Scotland and England  
 184 have most positive discussion toward performance trend. 39% of tweets in Scotland and 28% in England  
 185 has positive opinion about the performance of 5G.

**Table 4.** Comparison of combination of N-gram Features

Feature	Classifier	Accuracy	Precision	Recall	F-score
Uni + Bi	MLP	86.71	0.87	0.87	0.87
Uni + Bi	LR	86.14	0.86	0.86	0.86
Uni + Bi	Linear SVM	85.92	0.86	0.86	0.86
Uni + Bi	RBF SVM	62.61	0.72	0.63	0.54
Uni + Bi	NB	85.81	0.86	0.86	0.86
Uni + Tri	MLP	85.13	0.85	0.85	0.85
Uni + Tri	LR	86.59	0.87	0.87	0.87
Uni + Tri	Linear SVM	86.48	0.87	0.86	0.87
Uni + Tri	RBF SVM	62.61	0.72	0.63	0.54
Uni + Tri	NB	85.47	0.85	0.85	0.85
Bi + Tri	MLP	76.91	0.79	0.77	0.76
Bi + Tri	LR	77.02	0.79	0.77	0.76
Bi + Tri	Linear SVM	77.02	0.8	0.77	0.76
Bi + Tri	RBF SVM	62.01	0.6	0.59	0.6
Bi + Tri	NB	71.28	0.7	0.69	0.7

186 Figure 3 displays the negative sentiment towards fifth generation of cellular networks (5G) technology,  
 187 the radiation, price and below technology towards 5G. For example, most users concerns towards 5G signal  
 188 which it might cause cancer, the price of the 5G mobile which is very expensive and also the users concerns  
 189 about the performance of 5G mobiles. In order to find the most negative keywords towards 5G the word  
 190 frequency of the words negative sentences is calculated. The word frequency shows that the most discussed  
 191 keywords are radiation, price and performance. As Figure 3 (a) shows the most twitter users who located in  
 192 England, Scotland believe that the 5G can cause cancer. It is to be noted that, 31% of tweets in Scotland  
 193 and 31% in England has negative concern about the cause of radiation using 5G. In addition, the Figure  
 194 3 (b) most twitter users who located in Scotland believe that the price of 5G mobile is very expensive. It  
 195 is worth to mention that, 32% of tweets in Scotland and 27% of tweets in Wales has negative concerns  
 196 about the expensive price of 5G technology. Additionally, the Figure 3 (c) shows the most tweets users in  
 197 North-Ireland are less satisfied with current performance of the 5G mobiles. It is to be noted that, 27% in  
 198 North-Ireland and 25% in England has negative concern about current performance of 5G.

199 In contrast, the Figure 4 displays the overall trends for the twitter users. As the Figure 4 the most discussed  
 200 trend in agriculture (29%), healthcare (28%) and smart home (26%).

201 The Figure 5 displays the most discussed trends towards 5G technology, the radiation, price and below  
 202 technology towards 5G. For example, most users concerns towards 5G signal which it might cause  
 203 cancer, the price of the 5G mobile which is very expensive and also the users concerns about the low  
 204 technology of 5G mobiles. As shown in graph 5, England (a) the most discussed trends are agriculture  
 205 (31%), healthcare (28%) and smart-home(16%). In addition, Scotland (b) the most discussed trends are  
 206 agriculture (41%), healthcare (32%) and transport (12%). Additionally, in Wales the most discussed trends  
 207 are agriculture (29%), healthcare (28%) and transport(18%). Finally, North Ireland, the most discussed  
 208 trends are agriculture (29%), healthcare (28%) and transport (18%).

209 The graph 6 displays the occupation of the twitter users who had the most positive comments towards  
 210 fifth generation of cellular networks (5G).

211 The graph 7 displays the occupation of the twitter users who had the most negative comments towards  
 212 fifth generation of cellular networks (5G).





**Figure 2.** Positive trends towards fifth generation of cellular networks (5G)

213 As shown in Table 5, the top 10 positive and negative bigrams keywords related to next generation mobile  
 214 network (5G) are presented. As the Table III shows the tweets towards price of 5G network generation is  
 215 positive and towards EU countries are negative.

216 As shown in Table 6, the top 10 positive and negative trigram keywords related to next generation mobile  
 217 network (5G) are presented. As the Table IV shows the tweets towards price of 5G network generation is  
 218 positive and towards EU countries are negative.

## 5 DISCUSSION

219 In this study, the tweets related to the next generation mobile network (5G) were analysed and their  
 220 sentiment polarity were identified. The sentiment analysis of 5G keywords tweets considered because these  
 221 tweets carried lots of information related to next generation of communication network. Only English  
 222 tweets were considered because it has large coverage and the most widely common language used in the  
 223 world. In the current study, more than 10,000 tweets were retrieved. There were many duplicate tweets in  
 224 the retrieved data which were removed from the database. However, we did not limit our dataset to tweet



**Figure 3.** Negative trends towards fifth generation of cellular networks (5G)

225 discussion of online users, we used feed news related to 5G network communication to analyse because the  
 226 news consisted of valuable information related to incidents.

227 Moreover, as part of the data pre-processing we have removed the emoji characters from the tweets.  
 228 However, it is worth mentioning that emoji characters are commonly used in tweets and they are useful for  
 229 identifying the overall polarity of the tweets. In future work, we intend to consider emoji characters as part  
 230 of the process to identify the overall polarity of tweets as the emoji characters can provide more accurate  
 231 sentiment scores. In contrast, the time frame to analyse the specific tweets related to fifth generation of  
 232 cellular networks (5G) might be associated with events which increases the number of tweets in certain  
 233 dates, weeks or months. It is worth to mention that, there are lots of tweets related to conspiracy theories,  
 234 for example some of tweets believed that the 5G can cause covid-19 in people or it can cause cancer.

235 Herein we explain the most topics discussed which can change the overall sentiment polarity of the tweets  
 236 into positive:

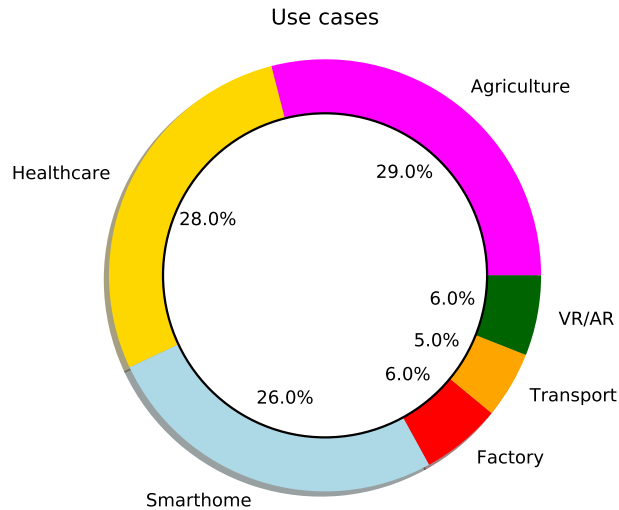


Figure 4. Trend for UK towards fifth generation of cellular networks (5G)

Table 5. Most frequent positive bigrams for next generation mobile network (5G)

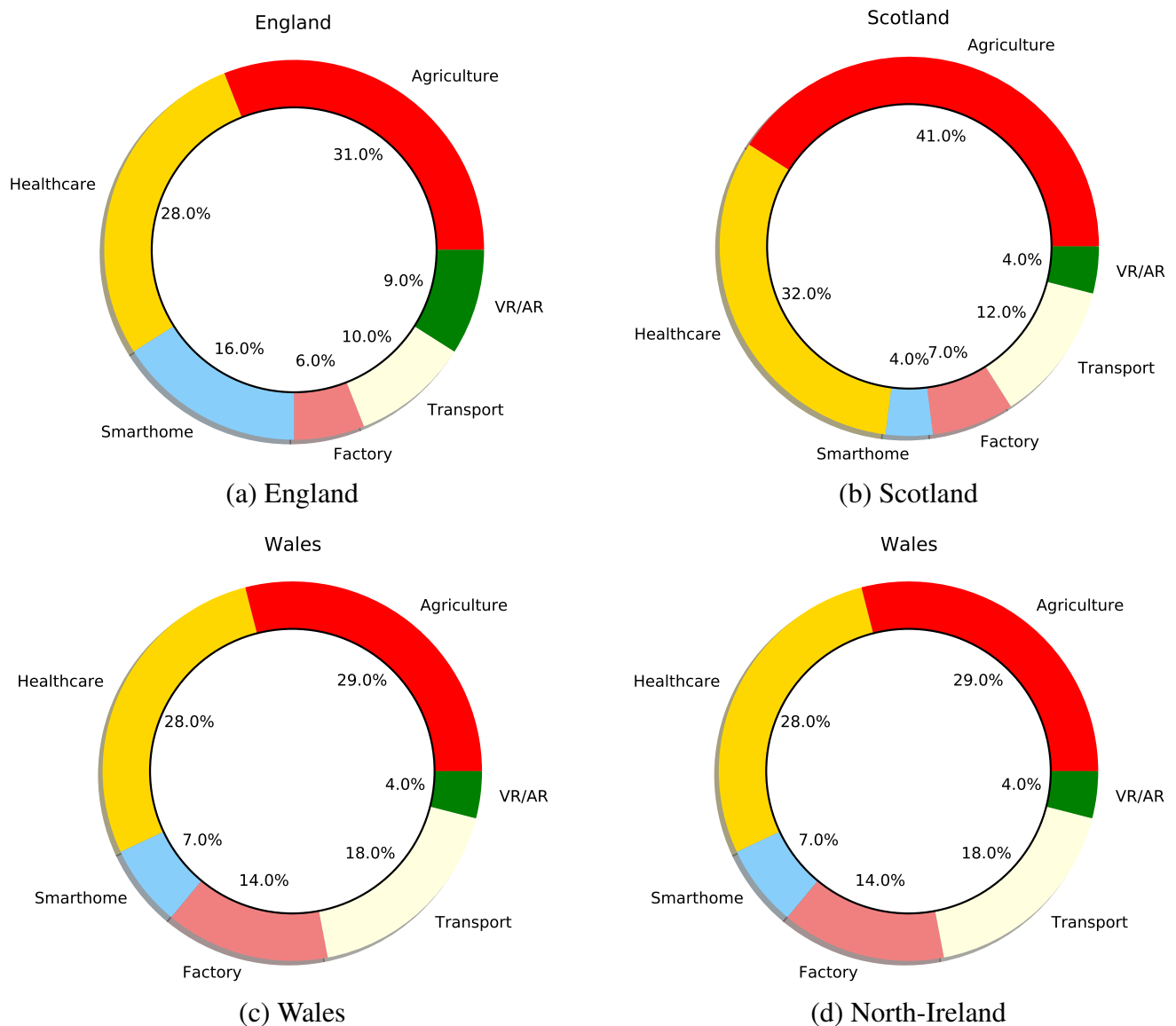
Positive Bigram	Negative Bigram
5G cheap	Low performance
Good coverage	Awful services
Good supply	blow Huawei
Great performance	Hate 5G
High security	Low speed
Fast speed	Low coverage
Great system	Low frequency
communication networks	5G Crap
Nokia performs	Slow 5G
great satisfaction	Expensive technology

Table 6. Most frequent positive Trigrams for next generation mobile network (5G)

Positive Trigram	Negative Trigram
cheap 5G phones	Slow coverage 5G
Good coverage phone	Low suppliers services
best 5G phones	blow Huawei China
Good Supply 5G	5G slow connection
Great areas coverage	5G low frequency
Good test zones	About 5G dangers
5G fast speed	Low frequency 5G
5G good communication	New 5G Crap
great 5G signal	Expect 5G Slow
Nokia performs well	Expensive New technology

237 **Speed:** The most discussed topics related to 5G was speed of the current network. Most of the tweets  
 238 were positive towards the speed of 5G. However, it is to be noted that, there are large number of tweets  
 239 were negative about the bandwidth and speech of the 5G mobile network.

240 **Security:** In addition, another most discussed topic related to 5G was security of the 5G services. Most  
 241 of the twitter users were positive towards the current security towards the services for 5G. For example,  
 242 “5G is more secure than 4G”.



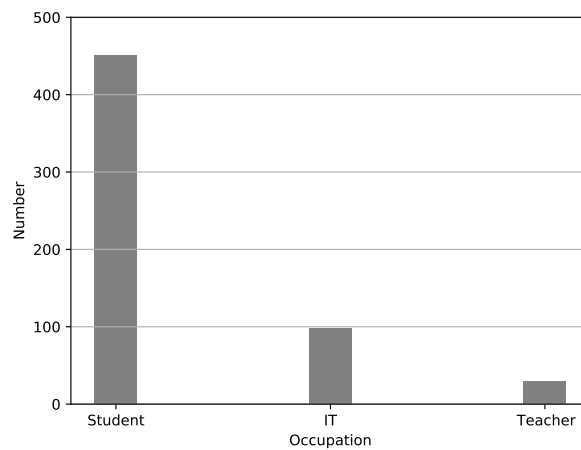
**Figure 5.** Most discussed trends for fifth generation of cellular networks (5G)

243 **Performance:** Additionally, the performance of the 5G mobile network were discussed in most of the  
 244 tweets. Most of the tweets were positive towards the performance of the 5G network communication. For  
 245 example, “I am really excited with current performance of 5G mobile”.

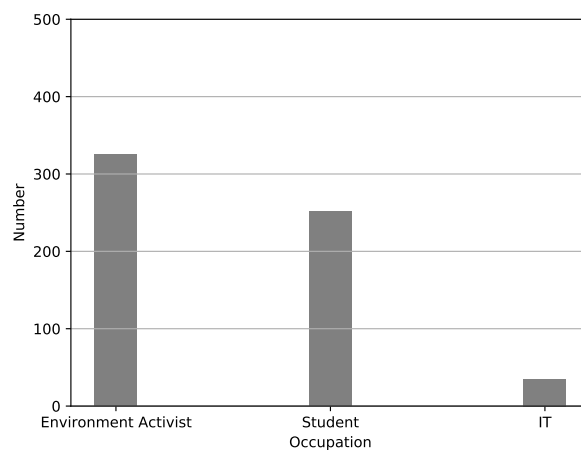
246 Herein we explain the most topics discussed which can change the overall sentiment polarity of the tweets  
 247 into negative:

248 **5G expose cancer:** Our finding shows that, most of the tweets were discussed about next generation  
 249 mobile network can cause cancer. For example, “I won’t buy 5G mobile, because it can cause cancer”.  
 250 There is strong negative comments towards the health issue towards 5G network communication. Most of  
 251 the Twitter users believe that the 5G can cause damage to their health.

252 **5G price:** Our analysis shows that, most of the Twitter users have negative comments towards the price  
 253 of the 5G mobiles. For example, one tweet clearly mentioned that “How can I afford to buy such expensive  
 254 mobiles”.



**Figure 6.** Occupation for most positive twitter users towards fifth generation of cellular networks (5G)



**Figure 7.** Occupation for most negative twitter users towards fifth generation of cellular networks (5G)

255 **Speed:** Furthermore, one of the most discussed topics were concern of twitter users about speed of the  
256 5G mobile network. For example, "I recently bought 5G mobile but I do not like the speed"

257 Table 7 shows the example of positive and negative tweets towards the fifth generation of cellular networks  
258 (5G).

**Table 7.** Example of Positive and Negative tweets towards Next Generation Mobile Network(5G)

<b>Positive</b>	<b>Negative</b>
Currently, The 5G phones are the best that you can buy right now	with 5G as dangerous to our health
5G has great performance	5G will be our deaths. Cancer rates will skyrocket.
I gotta say, the Vivo Nex 3 5G has a pretty great DAC.	5G is horrific, it causes Cancer, it's already banned in some countries.
The global economic potential of 5G is staggering It is predicted to add up to 3 million new jobs and create \$500 billion	it's not good enough the 5G privacy
I had good experience with 5G	5g is no good for humans health

## 6 CONCLUSION

259 In this study, we perform a series of sentiment analysis on data retrieved from twitter. The twitter data  
 260 under investigation was related to the fifth generation of cellular networks (5G). We have collected relevant  
 261 tweets in the English language. Therefore, we proposed framework for mobile networks (such as 5G)  
 262 based on different feature combinations. The performance of the proposed framework is evaluated using  
 263 different feature combination in terms of different evaluation metrics such as accuracy, precision, recall  
 264 and f-measure. In addition, we have compared the proposed method with different machine learning  
 265 algorithms such as Naïve Bayes, MLP, SVM, etc. In addition, we have analysed the tweets to understnad  
 266 user perception towards 5G. As part of our future work, we intend to extend the current framework for  
 267 multilingual sentiment analysis and integration of a closed loop self organising network algorithm with the  
 268 proposed user sentiment analysis framework.

## 7 ADDITIONAL REQUIREMENTS

269 For additional requirements for specific article types and further information please refer to Author  
 270 Guidelines.

## CONFLICT OF INTEREST STATEMENT

271 The authors declare that the research was conducted in the absence of any commercial or financial  
 272 relationships that could be construed as a potential conflict of interest.

## AUTHOR CONTRIBUTIONS

273 Author Contributions: Conceptualization, K.D., SH.A., Y.S., W.T., A.Z ,Q.H., M.A.I.; formal analysis,  
 274 K.D., SH.A., Y.S., Q.H.; investigation, W.T., SH.A.; resources, writing, review and editing, W.T., Q.H.A.,  
 275 M.A.I.; funding acquisition, Q.H.A., M.A.I.; All authors have read and agreed to the published version of  
 276 the manuscript.

## FUNDING

277 This research was funded under EPSRC DTA studentship (EPSRC DTG EP/N509668/1 Eng

## ACKNOWLEDGMENTS

278 This research work was funded by EPSRC DTG EP/N509668/1 Eng.

## REFERENCES

- 279 Adeel, A., Gogate, M., Farooq, S., Ieracitano, C., Dashtipour, K., Larijani, H., et al. (2019a). A survey on  
280 the role of wireless sensor networks and iot in disaster management. In *Geological disaster monitoring*  
281 *based on sensor networks* (Springer). 57–66
- 282 Adeel, A., Gogate, M., and Hussain, A. (2020). Contextual deep learning-based audio-visual switching for  
283 speech enhancement in real-world environments. *Information Fusion* 59, 163–170
- 284 Adeel, A., Gogate, M., Hussain, A., and Whitmer, W. M. (2019b). Lip-reading driven deep learning  
285 approach for speech enhancement. *IEEE Transactions on Emerging Topics in Computational Intelligence*
- 286 Ahmed, R., Gogate, M., Tahir, A., Dashtipour, K., Al-Tamimi, B., Hawalah, A., et al. (2021). Deep neural  
287 network-based contextual recognition of arabic handwritten scripts. *Entropy* 23, 340
- 288 Dashtipour, K., Gogate, M., Adeel, A., Algarafi, A., Howard, N., and Hussain, A. (2017a). Persian named  
289 entity recognition. In *2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive*  
290 *Computing (ICCI\* CC)* (IEEE), 79–83
- 291 Dashtipour, K., Gogate, M., Adeel, A., Hussain, A., Alqarafi, A., and Durrani, T. (2017b). A comparative  
292 study of persian sentiment analysis based on different feature combinations. In *International Conference*  
293 *in Communications, Signal Processing, and Systems* (Springer), 2288–2294
- 294 Dashtipour, K., Gogate, M., Adeel, A., Ieracitano, C., Larijani, H., and Hussain, A. (2018). Exploiting  
295 deep learning for persian sentiment analysis. In *International conference on brain inspired cognitive*  
296 *systems* (Springer), 597–604
- 297 Dashtipour, K., Gogate, M., Cambria, E., and Hussain, A. (2021). A novel context-aware multimodal  
298 framework for persian sentiment analysis. *arXiv preprint arXiv:2103.02636*
- 299 Dashtipour, K., Gogate, M., Li, J., Jiang, F., Kong, B., and Hussain, A. (2020). A hybrid persian  
300 sentiment analysis framework: Integrating dependency grammar based rules and deep neural networks.  
301 *Neurocomputing* 380, 1–10
- 302 Dashtipour, K., Hussain, A., and Gelbukh, A. (2017c). Adaptation of sentiment analysis techniques  
303 to persian language. In *International Conference on Computational Linguistics and Intelligent Text*  
304 *Processing* (Springer), 129–140
- 305 Dashtipour, K., Hussain, A., Zhou, Q., Gelbukh, A., Hawalah, A. Y., and Cambria, E. (2016a). Persent:  
306 a freely available persian sentiment lexicon. In *International Conference on Brain Inspired Cognitive*  
307 *Systems* (Springer), 310–320
- 308 Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Hawalah, A. Y., Gelbukh, A., et al. (2016b).  
309 Multilingual sentiment analysis: state of the art and independent comparison of techniques. *Cognitive*  
310 *computation* 8, 757–771
- 311 Dashtipour, K., Raza, A., Gelbukh, A., Zhang, R., Cambria, E., and Hussain, A. (2019). Persent 2.0:  
312 Persian sentiment lexicon enriched with domain-specific words. In *International Conference on Brain*  
313 *Inspired Cognitive Systems* (Springer), 497–509
- 314 Duong, T. Q., Chu, X., and Suraweera, H. A. (2019). *Ultra-dense networks for 5G and beyond: modelling,*  
315 *analysis, and applications* (John Wiley & Sons)
- 316 Gogate, M., Adeel, A., Dashtipour, K., Derleth, P., and Hussain, A. (2019a). Av speech enhancement  
317 challenge using a real noisy corpus. *arXiv preprint arXiv:1910.00424*
- 318 Gogate, M., Adeel, A., and Hussain, A. (2017a). Deep learning driven multimodal fusion for automated  
319 deception detection. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)* (IEEE),  
320 1–6
- 321 Gogate, M., Adeel, A., and Hussain, A. (2017b). A novel brain-inspired compression-based optimised  
322 multimodal fusion for emotion recognition. In *2017 IEEE Symposium Series on Computational*

- 323 *Intelligence (SSCI) (IEEE)*, 1–7
- 324 Gogate, M., Adeel, A., Marxer, R., Barker, J., and Hussain, A. (2018). Dnn driven speaker independent  
325 audio-visual mask estimation for speech separation. *arXiv preprint arXiv:1808.00060*
- 326 Gogate, M., Dashtipour, K., Adeel, A., and Hussain, A. (2020a). Cochleanet: A robust language-  
327 independent audio-visual model for real-time speech enhancement. *Information Fusion* 63, 273–285
- 328 Gogate, M., Dashtipour, K., Bell, P., and Hussain, A. (2020b). Deep neural network driven binaural audio  
329 visual speech separation. In *2020 International Joint Conference on Neural Networks (IJCNN) (IEEE)*,  
330 1–7
- 331 Gogate, M., Hussain, A., and Huang, K. (2019b). Random features and random neurons for brain-inspired  
332 big data analytics. In *2019 International Conference on Data Mining Workshops (ICDMW) (IEEE)*,  
333 522–529
- 334 Guellil, I., Adeel, A., Azouaou, F., Benali, F., Hachani, A.-E., Dashtipour, K., et al. (2021). A semi-  
335 supervised approach for sentiment analysis of arab (ic+ izi) messages: Application to the algerian dialect.  
336 *SN Computer Science* 2, 1–18
- 337 Hussain, A., Tahir, A., Hussain, Z., Sheikh, Z., Gogate, M., Dashtipour, K., et al. (2021). Artificial  
338 intelligence-enabled analysis of public attitudes on facebook and twitter toward covid-19 vaccines in the  
339 united kingdom and the united states: Observational study. *Journal of Medical Internet Research* 23,  
340 e26627
- 341 Jiang, F., Dashtipour, K., and Hussain, A. (2019). A survey on deep learning for the routing layer of  
342 computer network. In *2019 UK/China Emerging Technologies (UCET) (IEEE)*, 1–4
- 343 Jiang, F., Kong, B., Li, J., Dashtipour, K., and Gogate, M. (2020). Robust visual saliency optimization  
344 based on bidirectional markov chains. *Cognitive Computation* , 1–12
- 345 Kaity, M. and Balakrishnan, V. (2019). An automatic non-english sentiment lexicon builder using  
346 unannotated corpus. *The Journal of Supercomputing* 75, 2243–2268
- 347 Kumar, A. and Jaiswal, A. (2020). Systematic literature review of sentiment analysis on twitter using soft  
348 computing techniques. *Concurrency and Computation: Practice and Experience* 32, e5107
- 349 Kumari, P. and Haider, M. T. U. (2020). Sentiment analysis on aadhaar for twitter data—a hybrid  
350 classification approach. In *Proceeding of International Conference on Computational Science and*  
351 *Applications (Springer)*, 309–318
- 352 Liaqat, S., Dashtipour, K., Arshad, K., Assaleh, K., and Ramzan, N. (2021). A hybrid posture detection  
353 framework: Integrating machine learning and deep neural networks. *IEEE Sensors Journal* 21, 9515–  
354 9522
- 355 Liaqat, S., Dashtipour, K., Arshad, K., and Ramzan, N. (2020). Non invasive skin hydration level detection  
356 using machine learning. *Electronics* 9, 1086
- 357 Mamgain, N., Mehta, E., Mittal, A., and Bhatt, G. (2016). Sentiment analysis of top colleges in india  
358 using twitter data. In *2016 International Conference on Computational Techniques in Information and*  
359 *Communication Technologies (ICCTICT) (IEEE)*, 525–530
- 360 Mehta, R. P., Sanghvi, M. A., Shah, D. K., and Singh, A. (2020). Sentiment analysis of tweets using  
361 supervised learning algorithms. In *First International Conference on Sustainable Technologies for*  
362 *Computational Intelligence (Springer)*, 323–338
- 363 OPINCARIU, M. et al. (2019). Education in the 5g and the ai context. *Educația Plus* 23, 97–103
- 364 Ozturk, M., Gogate, M., Onireti, O., Adeel, A., Hussain, A., and Imran, M. A. (2019). A novel deep learning  
365 driven, low-cost mobility prediction approach for 5g cellular networks: The case of the control/data  
366 separation architecture (cdsa). *Neurocomputing* 358, 479–489



- 367 Sharma, A., Subramaniam, S. D., Ramachandran, K., Lakshmikanthan, C., Krishna, S., and  
368 Sundaramoorthy, S. K. (2016). Smartphone-based fundus camera device (mii ret cam) and technique  
369 with ability to image peripheral retina. *European journal of ophthalmology* 26, 142–144
- 370 Sharma, S. K., Daga, M., and Gemini, B. (2020). Twitter sentiment analysis for brand reputation of smart  
371 phone companies in india. In *Proceedings of ICETIT 2019* (Springer). 841–852
- 372 Shiva, A. S., Gogate, M., Howard, N., Graham, B., and Hussain, A. (2017). Complex-valued computational  
373 model of hippocampal ca3 recurrent collaterals. In *2017 IEEE 16th International Conference on*  
374 *Cognitive Informatics & Cognitive Computing (ICCI\* CC)* (IEEE), 161–166
- 375 Somula, R., Kumar, K. D., Aravindharamanan, S., and Govinda, K. (2020). Twitter sentiment analysis  
376 based on us presidential election 2016. In *Smart Intelligent Computing and Applications* (Springer).  
377 363–373
- 378 Taylor, W., Shah, S. A., Dashtipour, K., Zahid, A., Abbasi, Q. H., and Imran, M. A. (2020). An intelligent  
379 non-invasive real-time human activity recognition system for next-generation healthcare. *Sensors* 20,  
380 2653
- 381 Usama, M., Xiao, W., Ahmad, B., Wan, J., Hassan, M. M., and Alelaiwi, A. (2019). Deep learning based  
382 weighted feature fusion approach for sentiment analysis. *IEEE Access* 7, 140252–140260
- 383 Yadav, A. and Vishwakarma, D. K. (2020). Sentiment analysis using deep learning architectures: a review.  
384 *Artificial Intelligence Review* 53, 4335–4385