A Multi-layered Psychological-Based Reference Model for Citizen Need Assessment Using AI-Powered Models

Rajwa Alharthi^{1,2} · Abdulmotaleb El Saddik¹

Received: 30 April 2020 / Accepted: 26 July 2020 / Published online: 2 September 2020
© The Author(s) 2020

Abstract

We propose an automatic, low-cost, large-scale, nonintrusive human need recognition framework that utilized a multi-layered psychological-based reference model and designed with different modules including data collection, preprocessing, feature extraction and contextualization module. The reference model comprises several classification and regression models to identify human psychological needs, measure their satisfaction levels, evaluate their surrounding environment around different life aspects during any subjective event or towards emerging topics at any time, and in any location, using their publicly available social media content. We evaluate the predictive powers of various textual, psychological, semantic, lexicon-based and Twitter-specific features. To provide benchmark results, we compare and evaluate the performance of diverse machine learning algorithms. Our results confirm the effectiveness of the developed reference model. The framework is used to recognize citizen needs in response to the New Zealand terror attacks which occurred on March 15th, 2019.

Keywords Smart cities · Citizen needs · Reference model · Machine learning · Affect-aware city · Critical event · Need satisfaction · Social media analytics

Introduction

Urban innovation and solutions driven by Information and Communication Technologies (ICT) have been progressively applied to enhance urban life in terms of economy, mobility, environment, people, living and governance. The realization of a true smart city vision is now closer than ever [1]. More and more, the number of applications and services that are adopting these technologies with the intention of improving the performance of urban services which will, in turn, enhance the quality of life of citizens is growing [2]. For these applications to be effective, a variety of sensors are needed to continuously collect near real-time data. Currently, urban planners in a smart city rely mostly on the data obtained from measurement equipment or physical sensors

“hard sensors” such as cameras, environmental sensors, implanted medical devices, or telematics systems in vehicles [3]. The data retrieved from the deployed sensors are inserted into a large computing platform and then aggregated to provide a unified view of the city. Authorities then reference these data in making informed decisions on the management of the city and its events. The data retrieved from hard sensors, however, do not directly reflect the fluid response of people regarding changes in their immediate surroundings at any given time. Simply, deploying up-to-date technologies into the city’s systems does not solely make it smart. Incorporating innovative technologies with mechanisms to capture the citizens’ needs and feedbacks is the ultimate endeavor for the city to be considered smart. According to [2], engaging citizens in city planning and the development process is considered to be paramount in the evolution of smart cities, and, therefore, should be one of the main objectives when considering how to proceed. Moreover, in [4], Coe et al. mentions that encouraging the citizens’ participation in the decision-making process could lead to more democratic communities. Conclusively, more attention to the interaction between humans and urban space is needed to drive a more effective and relevant decision-making process. Consequentially, the principle of utilizing the citizen

✉ Rajwa Alharthi
ralha081@uottawa.ca; raharthi@tu.edu.sa
Abdulmotaleb El Saddik
elsaddik@uottawa.ca

¹ Multimedia Communications Research Laboratory,
University of Ottawa, Ottawa, ON K1N 6N5, Canada

² Department of Computer Science, Taif University,
Taif 26571, Saudi Arabia

themselves as “soft sensors” should be vitally considered to provide a successful and efficient implementation of a smart city.

The city should be able to understand, interpret and adapt to the affective states of its citizens, referencing their emotions, moods, and personality traits. The interpretation of affective data can be utilized in urban planning as a complement to the use of traditional hard sensors in evaluating ongoing planning processes, supporting decision-making, and providing additional insights regarding the city’s inhabitants to create a big picture of the city’s live state, as stated by an affect-aware city [5].

As we review previous works, we realize that recognizing a population’s diverse affective states is practical and achievable when taking advantage of the wealth of social media content that is constantly changing in a dynamic fashion, based on local and global happenings. User-generated content (UGC) from multiple social media platforms comprised of a vast amount of personal data, including users’ daily thoughts, insights, evaluations, feelings, and emotions, expressed through their geographic and time-based textual status updates. UGC is considered a rich source of information that can be used to reveal individuals’ affective states, and, as a result, researchers have begun mining this massive resource of affect data for this purpose. For example, a large part of the existing research focuses on general dimension sentiment analysis and opinion mining, under three polarity categories: positive, negative and neutral [6, 7] using lexicon-based methods and machine learning algorithms. Some studies go beyond the general sentiment analysis, recognizing distinct and dimensional emotional categories [8, 9] with the help of psychological lexicons. Certain recent studies have also explored more distinguishable long-term affective states such as personality [10] and mood [11], using psychological psychometrics.

Despite the recent focus on analyzing “how” people feel in a city by determining their distinctive affective states (emotions, personality and mood), the exploration around the “why” behind these feelings, actions and behaviors has, to date, received very little attention.

Being aware of citizen’s needs, within a smart city, can provide the root explanations of their motivation underneath their upsets, confusions and complaints based on many human need theories (HNT) [12–14]. This awareness can be utilized by city planners and decision-makers to adapt the city’s regulations, services, and plans in such a way to minimize and eventually reduce struggles, conflicts, and violent reactions.

Accordingly, our objective is to illuminate the vast benefits of automatically recognizing citizen needs. Therefore, in this work, we propose a theoretical-based multilayered reference model to assess citizen needs during any event at any time, and in any location using their publicly available

social media content. We design and develop need classification and regression models, with each corresponding to one of the concepts that form the layered reference model, namely need content recognition (NCR) model, need type identification (NTI) model, need satisfaction level measurement (NSM) model, social context evaluation (SCE) model, and life aspect identification (LAI) model. For a more comprehensive and deeper analysis, we develop the frustrated need intensity estimator (FNIE) model and the satisfied need intensity estimator (SNIE) model to determine the intensity score of the satisfaction level.

The rest of this article is structured as follows. In the next section, we provide a brief overview of the literature concerning the identification of human needs, followed by which we introduce the proposed multi-layered psychological-based reference model. Also, we present a need recognition framework that utilized the reference model with different specific designed modules to analyze citizen needs. In the subsequent section, we describe the experiments conducted to evaluate the proposed reference model and present the obtained results. Before the concluding section, we use our developed framework to analyze citizen needs in response to the New Zealand terror attacks that occurred on March 15th, 2019. Our conclusion is drawn in the final section, with a summary of the main findings and some suggestions for future directions.

Related Work

Individual needs are typically assessed using approaches from psychological science [15]. Due to many limitations, the traditional approaches are considered inadequate for large-scale need recognition and analysis. They are very time consuming and impractical when analyzing individual needs frequently in an interactive way within a large group (i.e. community). Moreover, using conventional methods is limited to a small group of respondents, which will only reflect a small percentage of the entire population within a city or community. Most of the conventional assessment surveys are designed with respect to one specific life aspect (e.g. relationships, work, etc.), and cannot be used to reflect multiple life aspects [16] and [17]. Moreover, they cannot efficiently capture the dynamics and context that embodies need experiences. All of these limitations are considered to be barriers to analyzing millions of individual needs.

On the other hand, few attempts have been made to infer people need using social media content. They differ in their objective, theoretical background, dataset used and their methodology. Despite the recognized potential of social media platforms like Twitter as a data source for analyzing human needs, the few existing works have some drawbacks and limitations. We will explain each work in details

clarifying their limitations and shortcomings in both the theoretical background and the recognition method.

IBM Research group [18] identified individual needs based on consumer behavior using Ford's model with the aim of enhancing the quality of direct marketing and influencing purchasing behavior. The model includes 12 need categories that correlate and explain consumer behavior, namely structure, practicality, challenge, self-expression, excitement, curiosity, liberty, ideal, harmony, love, closeness and stability. In the survey, they asked participant to list names of products they would like to buy and write their need that matches the model. The list of products mentioned in the responses are used to collect six million tweets to construct the data set and built their need model. Relying on satisfiers (i.e. products) in identifying individual needs is not efficient. Based on human need theory [19], satisfiers referring to the ways people satisfy their needs are changeable dependent on gender, age and culture. Therefore, we cannot rely on satisfiers to predict the need states. Moreover, the underlying need theory which is the Maslow hierarchy of fundamental needs [20] that used to build Ford's model was restrained by its cultural and hierarchical limitations [21].

Consider another mental state modeling work that was proposed by Rashkin et al. [22]. This work endeavors to understand the mental states of the actors in a simple story based on an event they experience. They proposed a dataset consisting of 15,000 commonsense stories,¹ which were manually labelled with emotion and motivation categories, using crowdsourced workers from Amazon Mechanical Turk. They applied motivation categories inspired by Maslow and Reiss' motivation theories, and for emotions, they referenced Plutchik's basic emotions model. During the annotation process, they found that the annotators were not familiar with motivational theories and, therefore, found it challenging when assigning motivation categories to the dataset. A logistic regression model is trained on TF-IDF, Neural Process Network (NPN), Recurrent Entity Network (REN), Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) encoders. The best model performance achieved a 35.23 micro- F_{score} using their proposed story dataset and attained a score of 64.8 when trained on the open text explanation provided by annotators using TF-IDF schema. Their classifier was most effective at predicting Maslow's physiological needs and Reiss's food motives only because they both have clear indication from the text. Their approach is designed based on long, formal stories and considers the multiple interactions between story characters.

Ding et al. [23] attempt the classification of sentences in personal stories from web blogs based on seven human needs. Their work shows numerous shortcomings in both the

theoretical background and the classification method. Not only did they fail to provide a concrete theoretical guideline in defining the need categories, they also did not consult an expert prior to proposing their taxonomy of needs and performing their annotation process. They claim their work was inspired by two human need theories: Maslow's Hierarchy of Needs and Fundamental Human Needs theories; however, their proposed need categories do not show any connection to either of those need theories. Another flaw lies in the fact that they associate the concepts of basic human needs with desires, wishes and goals, which are considered separate concepts and unique entities in most psychological need theories. Moreover, within their methodology, the number of instances in their dataset (approximately 559) is not sufficient for a classification task involving more than nine classes. This may explain the poor performance (54.8 average F_{score}) reflected in their results. The limitations and the drawbacks of the existing works motivate us to develop our psychological-based reference model to identify and assess citizen needs to promote well-being and prevent conflict and violence.

Methodology

We propose a human need recognition framework to help authorities to automatically monitor and analyze a population's needs at any time, any event and any location in the hopes of preventing conflict and violence, and also, to assure quality of life. In addition, they could interpret how people perceive their surrounding environment to recognize and understand the grounds of a population's discontent and determine the appropriate actions. Figure 1 illustrates the proposed psychological need recognition framework. The proposed framework employed a multi-layered psychological-based reference model and designed with different modules including data collection, preprocessing, feature extraction and contextualization module to be utilized in a wide range of applications with diverse contexts. The multi-layered psychological-based reference model which is the core element of the framework with all the designed modules will be explained in detail in the following sections.

Psychological-Based Multi-layered Reference Model

The psychological-based multi-layered reference model presented in Fig. 2 is developed based on several psychological theories [24, 25]. To have a comprehensive need analysis, multiple dimensions (i.e. concepts) of an individual's basic needs are adopted to construct the layers of the reference model. Layer 1 is formed to recognize need content by verifying the experienced emotion states. Based on human need theories, needs are inner states that cannot be recognized

¹ <https://uwnlp.github.io/storycommonsense/>.

Fig. 1 The psychological need recognition framework and its possible applications

Fig. 2 Psychological-based multi-layered reference model

directly. However, the self-determination theory (SDT) [26] states that needs are strongly correlated with emotions and feelings. Therefore, layer 1 aims to recognize emotions experienced by individuals to indicate their implicit needs. Layer 1.1 is constructed to identify the type of psychological needs that are present underneath the emotional states. In this layer, we considered the basic psychological needs proposed by the Basic Needs Theory (BNT) which are Relatedness, Competence and Autonomy [27]. These needs are fundamental for all individuals, regardless of their gender, age, culture, ethnicity and religion [28]. This layer can be extended by integrating other type of human needs.

Layer 1.2 is formed to measure the satisfaction level of the identified need. Relying on the polarity of the experienced

emotions, we can indicate need satisfaction level. Moreover, within layer 1.2, we determine the need satisfaction and frustration intensity score to best understand the individual's experience. The intensity refers to the degree or strength of the emotions felt. Each emotional word used by individuals is associated with differing intensities. For instance, both depression and unhappiness belong to an emotion class which would be categorized as sadness; however, they qualify differently in their intensities, where depressed conveys a higher amount of sad emotion than unhappiness. Because we understand that human needs are directly related to an individual's underlying emotions, we scrutinize the magnitude of the emotions in identifying the intensity of the satisfaction and frustration level. In analyzing a tweet of a specific need satisfaction level, the goal is to determine the degree of need satisfaction or need frustration felt by individuals. The intensity level can be determined using qualifiable categories (i.e. low, moderate and high) or using real value scores ranging from 0 to 1. A score of 1 means highest satisfaction level, whereas a score of 0 means the lowest. The ability to automatically determine and measure the intensity of satisfaction levels can be beneficial in many need recognition applications. For example, an application that monitors citizen psychological well-being could focus on detecting a significant need frustration in someone's life, and subsequently recommend an appropriate coping and healing strategy to prevent critical mental health issues such as depression or suicidal thoughts [29]. Another instance could be the early detection of high need frustration level in public reaction during any critical events or crisis could aid in determining appropriate

and immediate action which could prevent possible conflict and violence that could arise amongst the chaos.

Layer 1.3 has been constructed to assess the quality of individual's surrounding environment or social context. The layer aims to evaluate the way different social contexts range from interpersonal relationships (e.g. family) to more general settings, and distal contexts (e.g. economic and political systems) affect the need experience whether it helps to satisfy or frustrate the fulfilment of citizen need. To promote psychological well-being [30] and to avoid violence, citizens require full support from their social contexts.

Since basic needs must be satisfied across all life domains [12], layer 1.4 in the reference model is formed to identify the aspect of life that involve the need experience. We specified the most important life domains across the citizen life including family, social relations, work, education, government, leisure, health, religion aspects and general evaluation [24, 27, 31].

Data Collection Module

The data collection module is responsible for retrieving the data of interest from different social media networks, in this work we used Twitter platform. The textual and visual content will be retrieved with its metadata (i.e., posting time and geographical location) using Twitter Search API.² The social interaction information (i.e., retweets, replies and favourites) is also retrieved for each post. The textual data will go through the Data PreProcessing Module and Feature Extraction Module. After automatically recognizing citizen needs using the psychological-based reference model, the social interaction data and the metadata will be used by the contextualization module to represent the results.

Data Preprocessing Module

Several preprocessing tasks required to be performed to clean and prepare the text for the feature extraction steps and to further improve the performance. Based on the goal of our reference model, we performed the following combinations of preprocessing tasks: (1) Twitter non-textual elements such as URLs, video, image, mentions are filtered out. (2) Emojis, emoticons, hashtags, symbols and punctuations were kept since they have important expressive communication roles. (3) All the tweet elements are tokenized using TweetNLP Ark tool,³ a Twitter-specific tokenizer. (4) All tokens (i.e. word) are converted to lowercase and stemmed using Snow-Ball stemming algorithms.⁴

² <https://developer.twitter.com/en/docs/tweets/search/overview>.

³ <http://www.cs.cmu.edu/ark/TweetNLP/>.

⁴ <http://www.nltk.org/api/nltk.stem.html>.

Feature Extraction Module

Feature extraction is the process of representing each tweet as a feature vector. Each entry position in a vector corresponds to a feature type extracted from a tweet and represents the weight of that feature. We explore the usage of textual features, psychological features, semantic features, lexicon-based features and Twitter-specific features in recognizing citizen need. Table 1 shows the features used to develop the need models in each layer of the psychological-based multi-layered reference model.

Textual and Linguistic Features

Bag of Word Model (BoW) We used the well-known BoW model as textual features. The model considers all the distinct tokens (i.e. words, emojis and punctuation) in the dataset regardless of their order or the semantic dependency between them. We used Term Frequency-Inverse Document Frequency (TF-IDF) weighting schema to construct the feature vector for each tweet.

N-Gram Language Model (LM) We explore the continuous sequence of tokens with different n -sizes using N -gram models. LMs are useful for our identification tasks because they can capture more information. For example, it can capture some phrases and patterns (e.g. I like, I'm lonely), which the bag-of-words approach ignores. We extracted each n -gram token by calculating its TF-IDF weight.

Psycho-linguistic Features

We explore the use of different psychological features and evaluate how effective they are in providing insight into the way people express their needs on Twitter.

Linguistic Inquiry and Word Count (LIWC) We explore the LIWC collection of lexicons that developed based on psychology and cognitive theories [32]. LIWC lexicon consists of 4 dimensions (e.g. linguistic, psychological processes, personal concerns and spoken categories) and 92 lexicons which were developed by psychologists. It has been used to recognize different psychological concern and mental health disorders. Using LIWC, we extracted 92 features by calculating the percentage of total words in a tweet that match each of the predefined lexicons and formed the vector.

Linguistic Category Model (LCM) To analyze the language used in interpersonal events, we used LCM model, a psycho-linguistic classification approach, that classifies verbs people use during any social events [33, 34]. Using this approach, we can capture what is happening with a person during the event, his/her psychological state as well as the

Table 1 Features used to develop the need model in each layer of the psychological-based multi-layered reference model

Features	Layer 1	Layer 1.1	Layer 1.2	Intensity score	Layer 1.3	Layer 1.4
<i>Textual and linguistic features</i>						
Bag of words model (BoW)	x	x	x	x	x	x
N-gram language model (LM)	x	x	x	x	x	x
Part of speech (POS)					x	
<i>Psychological features</i>						
Linguistic inquiry and word count (LIWC)	x	x	x	x	x	x
Linguistic category model (LCM)		x			x	
<i>Semantic features</i>						
Emotion word embeddings (EWE)	x				x	
<i>Twitter-specific features</i>						
Emojis						
Emoji frequency	x	x	x		x	
Sentiment of emojis			x	x	x	
Categories of emojis		x	x			
The color of emojis		x	x			
Number of hashtags	x	x				
<i>Lexicons-based features</i>						
Sentiment lexicons						
NRC hashtag sentiment			x		x	
Negation lexicon					x	
MPQA					x	
Bing Liu opinion lexicon			x		x	
SentiWordNet				x		
Sentiment140				x		
SentiStrength (SS)				x		
AFINN				x		
SenticNet				x		
Emotion lexicons						
NRC word–emotion association lexicon (EmoLex)					x	
NRC hashtag emotion lexicon				x	x	
NRC affect intensity lexicon				x		
Life aspect lexicons						x

characteristics of the others involved in the event. The model consists of three linguistic categories: Descriptive Action Verbs (DAVs) that provide specific and concrete description of action during a short duration (e.g. hold, jog), State Verbs (SV) that describe thoughts (e.g. think, understand) and affective states (e.g. hate, respect) and Interpretative Action Verbs (IAV) that describe enduring behaviors and events without describing the feature of the action, such as (e.g. avoid, help, and attend). For the LCM features, we calculate the frequency of DAV, IAV and SV verbs in a given tweet.

Twitter-Specific Features

Due to the length limit imposed on some social media platforms such as Twitter, users tend to use emojis, emoticons, hashtags (#), mentions (@) and expressive symbols [35].

Thus, we adopt the Twitter-specific features including emojis and hashtags. We explore emoji in four ways: (1) the emoji frequency in a tweet, (2) the categories of emojis (e.g. foods, drinks, animal and nature) using Twitter emojis listed in Emojipedia,⁵ (3) the sentiment of the emojis using the sentiment lexicon created by Novak [36] and (4) the color of the emojis (e.g. black, cream white and dark brown). For hashtag features, we calculate the frequency of “#words” included in the text.

Lexicon-Based Features

Relying on the fact that the expressed emotions indicate the fulfillment of the needs, we adapt features that are driven

⁵ <https://emojipedia.org>.

by the use of the following existing sentiment and emotion lexicons to help measuring the satisfaction level in layer 1.2 [26, 37].

Bing Liu's Opinion Lexicon The Opinion Lexicon is a manually constructed lexicon from customer reviews about product's features [38]. It consists of 2006 positive words and 4781 negative words. For each tweet, we calculate the frequency of each word in the lexicons to form the feature vector.

Negation Lexicon Negation words can easily change the sentiment orientation from positive to negative, and in turn, completely reverse the intended meaning of the sentence [39]. Therefore, we used the count of the total occurrence of negated terms (cues or negation signals) such as don't, aren't, neither and never as features.

NRC Hashtag Sentiment Lexicon The NRC Hashtag Sentiment Lexicon⁶ is a large, automatically generated lexicon originated using 775,000 tweets that were retrieved using 78 unambiguous and strong positive and negative seeds of hashtags such as #amazing, #good, #excellent, #bad, and #terrible [40]. The lexicon contains 54,129 unigrams (single word), 316,531 bigrams and 308,808 pairs. Each of the words and the phrases is classified with a real-value score between $-\infty$ (most negative) and $+\infty$ (most positive). We have modified the NRC Hashtag Sentiment lexicon by excluding a number of elements (i.e. numbers, mentions, punctuation and other non-functional words) which we have deemed as useless in our scenario. For each tweet, we calculate the frequency of positive and negative words to form the vector.

Multi-perspective Question Answering (MPQA) Lexicon MPQA⁷ is a subjectivity lexicon that identifies subjectivity clues and aspects, including source of opinion, event, and sentiment expressions [41]. It contains a list of over 8000 subjective expressions collected from several resources. The expressions are manually compiled with prior polarities and tagged with their strength. For each tweet, we extracted two polarity features: positive words and negative words.

NRC Word–Emotion Association Lexicon NRC Word–Emotion Association Lexicon, also known as EmoLex, was constructed manually by conducting a tagging process using the Amazon Mechanical Turk crowdsourcing platform [42]. The Hashtag Emotion Corpus (aka Twitter Emotion Corpus, or TEC) was used to create the lexicon. The lexicon contains

14,182 unigram words tagged with eight emotion categories derived from the Plutchik's wheel of emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust). The words are also tagged with polarity (negative and positive). For each tweet, we extracted 8 features corresponding to the eight emotion category lists. We used EmoLex version 0.92. We also used the Expanded NRC Word–Emotion Association Lexicon, which expands the NRC word emotion association lexicon for the language used in Twitter.

Score-Based Sentiment and Emotion Lexicons

We used features derived from several sentiment and emotion-based lexicons. Specifically, we used score-based sentiment and emotion lexicons as features to determine the intensity score for satisfied and frustration needs. The intensity score of each employed resource was mapped at the same range from 0 to 1. The intensity score of each tweet is calculated by aggregating the intensity scores of each sentiment and emotional token (words and emojis) within the tweet provided by each of lexicons. For any given tweet, each token is reviewed to find a match in the lexicon. If a match is found, then the associated intensity score is retained. All individual intensity scores are compiled to calculate the overall score for each lexicon and use them as features. We use a variety of popular and comprehensive lexicons that differ in their data type (i.e. general, Twitter-specific lexicons, emotion based and sentiment-based lexicons), and in their annotation method (i.e. manually and automatically annotated lexicons). All the score-based sentiment and emotion lexicons are explained in detail below.

SentiWordNet The SentiWordNet lexicon⁸ is a publicly available synset-based sentiment lexicon introduced by Baccianella et al. [43]. It was constructed based on the WordNet lexical database synset, which is a set of synonyms meant to provide broad coverage. Each of the 115,000+ WordNet synsets is tagged with sentiment information automatically using semi-supervised machine learning method. Each synset is assigned a sentiment score value range between the interval [0.0 and 1.0] corresponding to its degree of positivity, negativity or neutrality. We used SentiWordNet version 3.0.

SentiStrength (SS) SentiStrength⁹ is a lexicon-based method proposed in [43] for use in determining sentiment strength. The lexicon was developed intended for use for short and informal language on any social media platform. It consists of a combination of 2546 booster words, emoticons,

⁶ <https://saifmohammad.com/WebPages/AccessResource.htm>.

⁷ <https://mpqa.cs.pitt.edu/>.

⁸ <http://sentiwordnet.isti.cnr.it/>.

⁹ <http://sentistrength.wlv.ac.uk/>.

negations and intensifiers from LIWC lexicon. The words in the lexicon are annotated manually based on a corpus of 2600 MySpace comments. Each lexicon entry is assigned a score indicating the polarity and the strength of the sentiment. Two scales are used to evaluate the sentiments. The first ranges from 1 (not positive) to 5 (extremely positive) to indicate the strength of positive sentiment, and, the second, employs -1 (not negative) to -5 (overly negative) to indicate the strength of negative sentiment. We used SentiStrength version 2.0. For each tweet, we extract two types of SentiStrength features: the sum of the positive scores and the sum of the negative scores.

AFINN AFINN¹⁰ is a manually constructed lexicon based on Bradley and Lane's Affective Norms for English Words (ANEW) lexicon [44]. ANEW provides emotional ratings for a large number of English words according to the psychological reaction of a person; however, it did not consider the informal words and slang that are commonly used in social media platforms. Nielsen in [45] filled that gap by creating the AFINN lexicon, an updated version of the ANEW lexicon, which focuses on microblogging platform language. Over 3300 English words and phrases were rated based on sentiment strength valence using a numerical sentiment score range from -5 which indicate a very strong Negative Sentiment and $+5$ for a very strong Positive Sentiment. For each tweet, we calculated the sum of the positive scores and the sum of the negative scores of the tweet words that matched the lexicon words.

Sentiment140 Sentiment140¹¹ is an automatically generated lexicon from a corpus consisting of a collection of 1.6 million tweets retrieved using these noisy labels: positive and negative emoticons [46]. The lexicon contains words and phrases which include 62,468 unigrams, 677,698 bigrams and 480,010 pair. Each of the listed words and phrases are associated with real-valued sentiment score between $-\infty$ (most negative) and $+\infty$ (most positive) and 0 for neutral. To ascertain the intensity of a tweet, we extracted two types of features. For each tweet, we add up the positive scores and the negative scores for each word in a tweet, separately.

NRC Affect Intensity The NRC Affect Intensity Lexicon¹² consists of 6000 words annotated with four basic association emotion labels: anger, fear, joy, and sadness [47]. All the words in each emotion category are rated with real-value numerical scores, illustrating the intensity strength for that emotion. For instance, "sohappy" is labelled as "joy" emotion, with an intensity score of 0.86. From the NRC Affect-

Intensity we extracted four features. The individual scores for each word in the tweet matching the four emotion classes are summed.

SenticNet SenticNet¹³ is a semantic and affective-based lexical resource that provides concept-level sentiment analysis for more than 100,000 natural language concepts [48]. It captures latent information in terms of semantics and sentics. Sentics are the affect information expressed in terms of four affective dimensions (pleasantness, attention, sensitivity, and aptitude). Also, SenticNet provides polarity-based scores between -1 (extreme negative) and $+1$ (extreme positivity). The intensity score is defined based on the sixteen basic emotions of the well-known Hourglass of Emotions. SenticNet does not rely only on the expressions that explicitly convey emotions using keyword counts and word co-occurrence frequencies; it is also able to leverage the implicit sentiments by analyzing expressions with semantically related concepts.

NRC Hashtag Emotion Lexicon Also known as NRC-Hash-Emo, an automatically created lexicon amassed from tweets with emotion word hashtags, utilizing The Hashtag Emotion Corpus (aka Twitter Emotion Corpus, or TEC) [41]. The corpus was constructed automatically using emotions hashtags such as #happy and #anger. The lexicon contains 16,862 emotion-word associations, including eight emotions (i.e. anger, fear, anticipation, trust, surprise, sadness, joy, and disgust). Each word in the lexicon is rated with real-value score between 0 (not associated) and ∞ (maximally associated).

Part of Speech (POS) POS tagging is the process of tagging a word with its part of speech such as verb, noun, adjective, or adverb based on the detected context. This process has proven to be effective in affects classification. To use POS tags as features, we used the CMU ARK Twitter Part-of-Speech Tagger.¹⁴ This POS tagger, developed specifically to be used with social media content, employs 25 tags that consider informal language and Twitter-specific properties [49]. Each token in a tweet is tagged with its POS. We calculated the frequency of each POS tag in a tweet to construct the feature vector. We have 25 different POS tags considered as features.

Semantic Features

Emotion Word Embeddings (EWE) Word embedding is a learning approach used to provide vector representations of

¹⁰ <https://darenr.github.io/afinn/>.

¹¹ <http://help.sentiment140.com/for-students/>.

¹² <https://saifmohammad.com/WebPages/AffectIntensity.htm>.

¹³ <https://sentic.net/>.

¹⁴ <http://www.cs.cmu.edu/ark/TweetNLP/>.

a particular word. It captures generic aspects of language structure, namely semantic and syntactic similarity between words. Words which frequently occur in similar contexts tend to be semantically similar and have similar regions of the vector space. Using word embeddings as features has recently proven to be effective in many text analysis tasks [50]. Pre-trained word embeddings are utilized when a text analysis task suffers from low resources and a large data set cannot be obtained [51]. It is very beneficial to incorporate this type of feature in designing our need models. In addition to capturing the semantic and syntactic context similarity, it is imperative in our need detection framework to preserve the emotional word meaning as well. Therefore, we used the pre-trained Emotion Word Embeddings (EWE) proposed by Agrawal et al. [52]. EWE is an emotion-enriched word embedding that enhances the regular embedding with respect to emotion analysis. This word embedding captures the affect information in tweets, while other existing generic word embeddings such as Word2vec¹⁵ and GloVe [53] capture only syntactic and semantic information. As mentioned in [52], according to the cosine similarity scores between the word vectors from the most popular pre-trained word embeddings (Glove and Word2vec), the word pair (happy, sad) is more similar than (happy, joy). This example highlights the limitation of using generic word embeddings in emotion-based analysis tasks that require more attention to affect expressions. Agrawal et al. [52] designed their emotion word embedding using an emotion model firmly grounded in psychology. By incorporating EWE as features, we can capture words with similar emotional meaning such as (happy, joy). The feature vector level within a tweet is calculated by aggregating the embedding values of the words within the tweet using an average word embedding scheme. We ended up with a dimensional vector of 400 for each tweet. Zero values are added for words with no corresponding embedding.

Life Aspect Lexicons Classifying a short text into one of nine life aspect categories can be very challenging, especially when we encounter an unbalanced dataset. External knowledge resources are necessary to help overcome this complication. Relying on lexical resources as external knowledge has proven to be beneficial during short text categorization and word sense disambiguation [54]. We constructed nine life aspect lexicons to help classify the input textual data in the most dominant life aspect. The construction of the nine life aspect lexicons involved the two following steps:

1. Selecting the initial seeds of words.

To establish an initial list of words related to each of the nine concepts of life aspects, we referenced online

dictionaries and thesauruses. One of such resources, Oxford mini-dictionaries, provides lists of words and phrases categorized based on different subjects and topics. We collected terms listed in each topic category that were relevant to an individual life aspect category. For example, the Oxford mini-dictionaries related to health domain consist of sub-categories including “diet”, “fitness”, “illness”, “medicine” and “mental health”. Furthermore, each of these sub-categories (e.g. illness) has their own sub-sub categories (i.e. “ailments and diseases”, “being ill”, “Injuries” and “Recovering from illness”). Each of the sub-sub-categories consists of lists of related words and terms such as “recover”, “healing” and “well”. This collection step was performed for all of the nine life aspect categories: education, family, work, health, government/political, social relation, leisure, religion, and general evaluation. For the general evaluation aspect, we constructed the lexicon using words and phrases that express consistent and constant occurrences.

2. Expanding the words lists.

Once the initial list of words was amassed, we moved to the second stage where we extend the list of words in each lexicon by identifying and collecting synonyms and alternative expressions. WordNet, a large and well-constructed database, provided a good starting point. WordNet is a lexical database developed at the Cognitive Science Laboratory at Princeton University.¹⁶ WordNet organizes the knowledge in a way that reflects current psycholinguistic theories regarding how humans archive their lexical memories. The information in WordNet is catalogued into logical groupings based on sets of cognitive synonyms called synsets. Each synset consists of a list of synonymous words that are inter-linked by means of conceptual semantic similarity relations. Synsets are built with an intrinsic correlation between a concept and its corresponding words which are interchangeable among many contexts. The semantic relations have different types, including (Is-a/Has-a) relation, (Part-of/Has-part) relation, (Member-of/Has-member) relation and (Substance-of/Has-substance) relation. Figure 3 shows a graph from the WordNet visual dictionary¹⁷ to explain (Is-a) relation. Table 2 shows examples of words included in each of our life aspect lexicons.

The life aspect features are extracted as follows: first, we count all the text elements in a tweet, including punctuations, numbers, emoji and words. Then we compare each word in a tweet against all words in the predefined life aspect lexicons. The percentage of each life aspect lexicon is calculated by dividing the number

¹⁵ <https://code.google.com/archive/p/word2vec/>.

¹⁶ <https://wordnet.princeton.edu/>.

¹⁷ <http://wordvis.com/>.

Fig. 3 A graph from the WordNet visual dictionary showing synonyms of the word “education” using (Is-a) relation

of its word frequency by the total number of the structure elements of the tweet, then multiplying that by 100. Similar phrases such as (all the time) and (every time) from a general life evaluation lexicon are counted as a single word rather than each separate word. Figure 4 explains how life aspect lexical features are extracted. In the example, the tweet has 21 elements. Four words from the work category lexicon compose 19.04% of the tweet, and one word from leisure category lexicon represents 4.761% of the tweet.

Designing and Developing Psychological Need Models

To design and develop the classification and regression psychological need models within the reference model’s

Fig. 4 Extracting life aspect features from a tweet

layers, we adopt the learning method explained in [55]. Based on the first 3 layers of the psychological-based reference model, authors designed and developed the need content recognition (NCR) model, the need type identification (NTI) model and the Measuring Need Satisfaction Level (MNS) model. In this paper, we extend the work by exploring different feature sets as well as different machine learning algorithms. Also, we design and develop the social context evaluation (SCE) model and the life aspect identification (LAI) model that are corresponding to layer 1.3 and layer 1.4, respectively. In the process of developing the SCE and the LAI psychological need models, we used the psychological need dataset proposed in [56]. In dataset we faced imbalanced problem for layer 1.3 and layer 1.4. Therefore, we used different sampling techniques including resampling and SMOTE to balance the classes’ distribution as Table 3 shows [57]. In layer 1.3, we used under-sampling technique for “Not clear” class to randomly remove 32.73% of the instances. For the minority class “Support social context”, we increased the instances by 120%. In layer 1.4, we decreased the amount of instances by 42.3% for the majority class “Social relation” and increased the minority classes “Religion” and “Government” by 400 and 300%, respectively, using SMOTE.

Moreover, we extend and improve layer 1.2 by developing the frustrated need intensity estimator (FNIE) model and the satisfied need intensity estimator (SNIE) model to determine the intensity score of the satisfaction level. To ascertain the intensity score of the need satisfaction level, we used an existing emotion intensity dataset called EmoInt to train and evaluate our proposed intensity estimator models. EmoInt is a benchmark tweet dataset with affect-related intensity score proposed for WASSA-2017

Table 2 Examples of words included in the life aspects lexicons

Life aspect category	Examples of words included in the lexicons
Work	Tasks, occupation, resignation, practicing, agents, executive, company, boss, investment, qualified
Education	Learn, diploma, exam, graduate, academic, grade, school, course, GPA, study
Health	Anxiety, blood, pressure, workout, painful, drug, surgery, relief, medical, stroke, healthcare
Social life	People, enemy, human, who, relationship, friend, boys, someone, women, guy
Government	Prime minster, military, freedom, vote, mayor, poll, reform, campaign, policy, election
Leisure	Play, travel, song, paint, game, fun, camp, music, festival, holidays
Religion	Church, God, belief, pray, Jesus, buddha, worship, mercy, Christian, Allah
Family	Grandma, father, married, spouse, mother, husband, son, boyfriend, girlfriend, sister
General evaluation	Every time, always, forever, until now, almost, constantly, continually, nonstop, 24/7, evermore

Table 3 Psychological need data set statistics before and after applying resampling

Layers	Class name	Resampling technique	Amount of resampling	Before resampling	After resampling
Layer 1.3	Supportive social context	SMOTE	120%	947	2083
	Not supportive social context	–	–	1935	–
	Not clear	Under-sampling	32.73%	3452	2322
Layer 1.4	Work	–	–	487	–
	Education	–	–	345	–
	Health	–	–	307	–
	Family	–	–	876	–
	General evaluation	–	–	396	–
	Government	SMOTE	300%	80	320
	Leisure	–	–	963	–
	Religion	SMOTE	400%	61	305
	Social relation	Under-sampling	42.3%	2196	1266
	Other	–	–	650	–

shared task on emotion intensity [58]. In the construction of the EmoInt dataset, 24 ranking judges manually annotated 7097 tweets based on four emotion categories: anger, fear, joy, and sadness, along with their associated intensity score range from 0 to 1. Since satisfaction in reference to individual need presents positive emotions and frustration needs lead to negative emotions, we divided the EmoInt dataset into positive and negative emotion datasets. The Negative intensity dataset included 5465 tweets which reflected the emotions anger, fear, and sadness, while the positive intensity dataset encompasses 1608 tweets of a happy and joyful emotion. For each of the datasets, we trained a regression model to determine the intensity score. In layer 1.2, if the tweet was classified to have a satisfied need expression, we applied the satisfied need intensity estimator (SNIE) model within the same layer to derive the intensity score of the satisfied need. In the case of tweets classified as frustrated, we used the frustrated need intensity estimator (FNIE) model to determine the frustration intensity score.

After extracting the features listed in Table 1 for each layer for each layer of the reference model, all the features are normalized using Min–Max scaling. Before training the machine learning algorithms, we removed irrelevant and noisy features using a dimensionality reduction technique GainRatio [59].

Contextualization Module

The post-interaction data (i.e. retweet, favourites and mentions) and metadata (i.e. time and location) will be used by the Contextualization Module for result analysis and representation.

Experiment and Evaluation

We experimentally evaluate the performance of our proposed reference models on the psychological need dataset and the EmoInt dataset. The experiments are designed with the goal of measuring the effectiveness of the five classification models of our framework: NCR, NTI, NSM, SCE and LAI models and the two intensity estimator regression models: FNIE and SNIE. Since we are reporting the result of baseline, a set of experiments were conducted to validate the importance of different textual features, semantic features, psychological features and Twitter-specific features along with different machine learning algorithms.

Experimental Setting

The psychological human need dataset was divided into 70% for training and 30% for testing, while preserving the class distribution. For the classification tasks, we explored the most well-known machine learning algorithms, which we singled out based on their various learning methods and techniques, to provide benchmark results. Based on the best practices for classical machine learning engineering, we have selected the Multinomial Naive Bayes algorithm (MNB) as a probabilistic classifier, the Support Vector Machine algorithm (SVM) as a linear-decision boundary algorithm, the Logistic Regression (LR) as a regression classifier, the Random Forest algorithm (RF) as an ensemble classifier, the K-Nearest Neighbors algorithm (K-NN) as an instance-based algorithm, and the Decision Tree algorithm (DT) as a rule-based classifier [60]. It is worth noting that in 2012, Hinton and Krizhevsky [61] and [62] showed that deep learning functions are typically superior to other machine learning

Fig. 5 The effect of different textual features on the accuracy of the NCR model

methods for large datasets where complex processing is required. Since we have a linear classification process as well as our dataset is comparably small, deep learning analysis and prediction are kept outside the main scope of this paper.

For regression tasks, we used the Support Vector Regression (SVR) algorithm. We split each of the positive and the negative intensity datasets into 70% for training and 30% for testing.

We used the machine learning algorithm libraries implemented in *Weka*. *Weka* stands for *Waikato Environment for Knowledge Analysis*, a machine learning software developed at the University of Waikato [63]. While conducting the experiments, we followed an incremental feature selection approach. We comparatively evaluated the predictive power of each distinct feature and reported the obtained result before and after any feature addition, both as a set and individually.

Evaluation Metrics

The classification models' performance is assessed using the metrics of recall, precision, F_{score} and accuracy. The regression models are evaluated using the Pearson's correlation coefficient (r) of model predictions with the EmoInt gold ratings and the root mean squared error (RMSE). RMSE calculation is used in indicating the average model prediction error between actual and predicted data. The usage of error square is important to statistically weigh any irregular values that might affect the results. Compared to traditional error calculations such as error rate (ER) and mean absolute

deviation (MAD), RMSE performs better in terms of error prediction accuracy as well as monotonicity since it has the advantage of penalizing large errors in predicting observations according to the variables of previous results.

Experimental Results

The Effectiveness of the Need Content Recognition (NCR) Model

We first study the impact of the different textual features BoW and n -gram models, including unigrams, bigrams and trigrams, on the model's performance. We consider how they interact both individually and in combination. The graphical comparison in Fig. 5 illustrates these interactions. As we can see from the graph, using unigrams and the combination of all the unigrams, bigrams and trigrams yields the best performance in comparison to the use of bigram and trigram models. The combination of the unigrams, bigrams and trigrams provides slightly better results when used for the following classifiers: SVM with an accuracy of 64.88%, K-NN 62.1%, RF with 67.18% and DT shows 66.5% accuracy. Next, we investigate the impact of the LIWC psychological lexicon, emojis and the number of hashtag features on the models' performance. Table 4 shows the accuracy and F_{score} for all the selected machine learning algorithms, using all the features both as a set and individually. Table 4 shows we notice that using the LIWC psychological lexicon alone has a positive influence on the need content recognition model for all classifiers. In addition, considering the frequency of use of emojis seems to provide a slight advantage to all the

Table 4 Accuracy and F_{score} of the NCR model, comparing the use of different features

Features	SVM		MNB		LR		K-NN		RF		DT	
	A	F	A	F	A	F	A	F	A	F	A	F
LIWC	67.05	0.65	60.83	0.51	65.94	0.64	58.44	0.58	65.69	0.64	59.45	0.59
Emotion word Embeddings (EWE)	66.74	0.65	58.91	0.54	66.31	0.65	59.61	0.59	64.39	0.61	57.26	0.57
LIWC + emojis	67.09	0.66	60.91	0.52	65.94	0.64	58.66	0.58	66.29	0.64	59.55	0.59
LIWC + emojis + hashtag numbers	66.35	0.652	60.89	0.52	65.67	0.64	58.44	0.58	65.67	0.64	59.59	0.59
LIWC + emojis + hashtag numbers + emotion word embeddings (EWE)	68.72	0.68	61.95	0.59	68.74	0.67	60.77	0.60	65.11	0.62	59.92	0.59
LIWC +3 $\geq$ gram +1 + emojis + hashtag numbers + emotion word embeddings (EWE)	64.81	0.68	72.67	0.72	70.54	0.70	62.00	0.61	65.91	0.63	61.23	0.61

models. There are no significant changes in the model accuracies after combining the LIWC lexicon with the frequency of Emoji and the number of hashtag features. Combining the LIWC lexicon, the frequency of Emoji, the number of hashtags and the n -gram model have improved the accuracy significantly for MNB classifier by 11.78% and slightly for LR classifier by 4.87%, K-NN by 3.56% and DT by 1.64% and no change in RF accuracy. SVM classifier achieved an accuracy of 64.81%, which is lower when compared with the accuracy achieved with individual features, due to the large number of features used.

The problem of data sparsity and high dimensionality is solved using the GainRatio. Out of the 24,013 features, we examine the most predictive features using different threshold values (0.01, 0.05 and 0.09). Using a threshold value of 0.01 leaves us with 2526 features, and improved the accuracy for SVM by 13.9%, RL by 0.64%, and DT by 5.51%. Using 0.05 threshold value resulted in 2106 features and boosted the accuracy by 6.32% for MNB, 13.29% for K-NN and 10.72% for RF, as shown in Fig. 6. The MNB classifier achieved the best accuracy in recognizing need content with at 79.13%, with a 0.77 F_{score} , as Table 10 shows.

The Effectiveness of the Need Type Identification (NTI) Model

The experiment results with textual features show that the combination of the unigrams, bigrams and trigrams achieved better results for SVM, MNB and RF classifiers, as depicted in Fig. 7. Using only the unigram feature provides the best accuracy for K-NN and DT, while the combination of unigrams and bigrams are best for RF classifier. In Table 5, the outcomes of using the other features are listed. All the classifiers achieved acceptable ranges of accuracies after applying the psychological features derived from the LIWC and LCM lexicons (DAVs, IAV, SV). This is achieved because LIWC lexicon has different psychological dimensions that can be directly linked to each need types. The addition of the number of hashtag features and the frequency of emoji has slightly increased the accuracy for SVM and RF. Using all the emoji features, including the frequency of emoji, the categorized emoji and the colored emoji have not improved the accuracies except for a slight increase within the SVM classifier. The combination of all the features results in the maximum accuracy for all the classifiers, especially for MNB classifier, with an absolute accuracy gain of 13.31%. After using GainRatio to select the best features out of all 9500 features for each classifier with different threshold values, the accuracy increases marginally, as shown in Table 10 and Fig. 8. The best classifier in identifying need type is SVM, showing an accuracy rate of 81.97% and a 0.81 F_{score} .

Fig. 6 NCR model accuracy before and after selecting the most predictive features using GainRatio

THE EFFECT OF TEXTUAL FEATURES ON (NTI) MODEL

Fig. 7 The effect of different textual features on the accuracies of the NTI model

The Effectiveness of the Need Satisfaction Level Measurement (NSM) Model

Based on the comparison graph for BoW and n -gram models in Fig. 9, the highest accuracy was achieved by SVM, LR and DT classifiers when combining unigrams, bigrams and trigrams, whereas when combining unigram and bigrams, MNB and RF achieved the best accuracies. K-NN classifier responded best when using the unigram model alone. As illustrated in Table 6, the RF and DT models performed very well and exhibited high accuracy rates when applying the LIWC psychological lexicon. Since LIWC lexicon has dimension for affective processes with many sentiment (positive and negative) and emotion categories (anxiety, anger and sadness), this could lead the increased in the performance. Adding emoji features, including the frequency of emoji, the sentiment, the categorized, and the colored emojis, increases the accuracy slightly, with the exception of the K-NN classifier. While adding the NRC and the opinion sentiment lexicons did not significantly affect the accuracy on most models, there was a slight rise (2.17%) in accuracy for the DT classifier. A markedly high performance is produced for all the classifiers except K-NN when combining all

the features, especially the SVM classifier, which achieved 93.10% the heist accuracy. The K-NN classifier archived its best result using the LIWC psychological features alone, effectuating an accuracy of 75.31%. In the feature selection technique, applying the most predictive features with a threshold of 0.01 gives the highest accuracy of 93.56 % when used for SVM, and 92.25% for MNB, as shown in Table 10 and Fig. 10. A threshold of 0.05 gives the best result for SVM classifiers, showing 93.56% accuracy, and also for RF (92.28% accuracy). For K-NN classifier, eliminating the features with a 0.09 threshold value increased the accuracy to its best (89.84%). Among all of the evaluated machine learning algorithms, SVM is the best in determining the need satisfaction level, Table 10.

The Effectiveness of the Intensity Estimator Models

Results show that using the combination of unigram and bigram models as well as the combination of all the n -gram models provides good outcomes r (0.62) for SNIE and r (0.61) for FNIE) and tend to be more predictive than the other models as Fig. 11 shows. As illustrated in Tables 7 and 8, using the LIWC lexicon leads to an average r of (0.58)

Table 5 Accuracy and F_{score} of the NTL model, comparing the use of different features

Features	SVM		MNB		LR		K-NN		RF		DT	
	A	F	A	F	A	F	A	F	A	F	A	F
LIWC	68.28	0.68	63.30	0.63	68.34	0.68	67.08	0.66	71.76	0.71	60.35	0.60
LIWC + LCM (DAVs, IAV, SV)	68.82	0.68	63.40	0.632	68.98	0.68	66.82	0.65	72.34	0.720	60.51	0.60
LIWC + emojis	68.66	0.67	63.09	0.629	68.61	0.68	66.92	0.65	72.39	0.72	59.77	0.59
LIWC + emojis + LCM (DAVs, IAV, SV)	68.76	0.682	63.30	0.63	69.08	0.69	67.19	0.66	73.02	0.72	60.30	0.60
LIWC + emojis + LCM (DAVs, IAV, SV) + hashtag numbers	69.24	0.69	63.51	0.63	69.13	0.69	66.92	0.65	72.45	0.72	59.56	0.59
LIWC + emojis + LCM (DAVs, IAV, SV) + hashtag number + categorized emojis	69.40	0.69	63.35	0.63	68.92	0.68	66.08	0.65	73.55	0.73	59.56	0.59
LIWC + emojis + LCM (DAVs, IAV, SV) + hashtag number + categorized emojis + colored emojis	70.03	0.70	63.03	0.62	68.71	0.68	65.98	0.65	72.45	0.72	59.41	0.59
LIWC + 3 $\geq$ gram ≥ 1 + emojis + LCM (DAVs, IAV, SV) + hashtag number + categorized emojis + colored emojis	79.91	0.79	76.34	0.76	72.92	0.73	70.82	0.70	77.65	0.77	66.50	0.66

for the SNIE model, and (0.51) for FNIE model, suggesting that the psychological I features are effective for intensity determination. Adding the NRC affect intensity lexicon led to a slight increase in r score for both models. Applying SentiStrength does not have a significant impact on the SNIE model; however, it did improve the performance of the FNIE model. Using the AFNN lexicon has the opposite effect, where a slight improvement was noted in the SNIE model's performance, but not for the FNIE model. The experiments also determined that there is improvement for both models when using S140 lexicon and Sentiment Emojis features. The addition of the Hashtag Sentiment Lexicon and the SentiWordNet lexicon did not positively affect the model performance due to the presence of redundant features in the lexicons. Among all the lexicons utilized, LIWC is the most predictive single lexicon. Combining all the score-based sentiment and emotion lexicons with the textual features produces a statistically significant improvement and has boosted the correlation r from (0.64) to (0.71) for SNIE and for FNIE from (0.59) to (0.68). As can be observed from Table 9, the FNIE model has a slight advantage over the SNIE model as its predicted result will only deviate 13 points on a scale of 100 from the actual value.

A noticeable improvement is also obtained after applying the GainRatio, which increases r values from (0.71) to (0.73) for SNIE model and (0.68) to (0.72) for FNIE model as Table 9 and Fig. 12 show.

The Effectiveness of the Social Context Evaluation (SCE) Model

As Fig. 13 shows, after comparatively evaluating the predictive power of the BOW model and the n -gram models, the combination of the unigram, bigram and trigram is selected because it provides the best results for most of the classifiers. As demonstrated in Table 11, using LIWC and LCM (DAVs, IAV, SV) psychological features also results in decent outcomes for all the classifiers. RF achieved the highest accuracy with 71.45%, and lowest was DT at 60.41%. Using the pre-trained emotion embeddings ended up with acceptable results for some classifiers, including SVM (65.45%), RF (65.82%), LR (64.87%) and K-NN (59.83%), while a lesser performance was noted for the MNB (48.63%) and DT (49.36%) classifiers. Using sentiment-based lexicon features showed satisfactory results among most of the classifiers; however, all the classifiers, except DT, achieved better accuracy when using emotion-based lexicon features. Specifying the emotion categories such as “disappointment” “anticipation” and ‘Trust’ helped to determine the surrounding social context type. For all classifiers, combining the psychological feature LIWC with emotion-based lexicon features enhanced the accuracy more than when these features are utilized individually. Combining the 1–3 g model with

Fig. 8 NTI model accuracy before and after selecting the most predictive features using GainRatio

THE EFFECT OF TEXTUAL FEATURES ON (NSM) MODEL

Fig. 9 The effect of different textual features on the accuracies of the NSM model

the psychological features, including LIWC and CML lexicon, boosted the accuracy for most of the classifiers: SVM by 5.25%, MNB by 8.09%, LR by 3.16%, RF with 3.52% and DT, 3.05 percentage points. Adding the sentiment and emotion-based lexicons, the frequency, and the sentiment of emoji to the previous features marginally increases the accuracy for MNB, LR and DT. MNB and LR achieved their best results by combining all the features, ending up with 72.34 and 69.50%, respectively. Experiment results showed that, in the case of using GainRatio with a 0.01 threshold, the accuracy was augmented for LR, K-NN and DT classifiers. At a 0.09 threshold value, the accuracy of MNB and SVM improved (Fig. 14). As Table 10 shows, RF achieved the best result in evaluating social context (surrounding environment) when using the LIWC and LCM (DAVs, IAV, SV) psychological features in combination with textual features (Table 11).

The Effectiveness of the Life Aspect Identification (LAI) Model

The experiment with textual features shows that the unigram model gives the best performance in identifying life

aspect when compared to the bigram and trigram models, and also, with the combination of the models. This indicates that when identifying life aspect, it is not necessary to detect phrases and sequence of words, as in the previous need layers. Individual terms are more effective. Figure 15 provides a graphical comparison between classifiers using textual features. As Table 12 shows, acceptable results were achieved for some classifiers using the LIWC lexicon, showing 51.69% for SVM, LR with 51.65%, and RF at 51.80%.

Meager results are noted for MNB with 39.06%, K-NN at 36.97%, and DT coming in at 39.79%. Adding the eight life aspect lexicons improved the accuracy for all the classifiers. The experiments show that combining the unigram model with the LIWC lexicon optimized the accuracy for LR at 56.82%, while combining the unigram model with the eight life aspect lexicons provided the best accuracy rates for MNB (57.72%), K-NN (45.26%) and RF (55.46%). For SVM and DT, combining all the features improved their accuracy to 55.52 and 49.43%. As Fig. 16 shows, eliminating the useless features as we measure a model's performance improved the accuracy rates for all the classifiers: SVM by 4.9%, MNB at 2.03%, LR by 4.25%, K-NN by 6.37%, RF by 1.75%, and

Table 6 Accuracy and F_{score} of the NSM model, comparing the use of different features

Features	SVM		MNB		LR		K-NN		RF		DT	
	A	F	A	F	A	F	A	F	A	F	A	F
LJWC	72.10	0.71	65.40	0.63	72.87	0.72	75.31	0.75	89.65	0.89	80.70	0.80
LJWC + emojis	71.21	0.70	65.36	0.63	72.99	0.728	75.125	0.74	89.57	0.89	80.70	0.80
LJWC + emojis + sentiment emojis	71.98	0.71	65.90	0.63	73.42	0.73	74.89	0.74	90.15	0.90	82.48	0.82
LJWC + emojis + sentiment emojis + categorized emojis	72.26	0.72	66.21	0.64	73.49	0.73	74.97	0.74	90.31	0.90	83.10	0.83
LJWC + emojis + sentiment emojis + categorized emojis + colored emojis	72.41	0.71	66.05	0.64	73.45	0.73	75.04	0.74	90.43	0.90	83.22	0.83
LJWC + NRC hashtag sentiment + opinion lexicons	71.52	0.71	65.71	0.63	73.11	0.72	75.28	0.75	90.12	0.90	81.05	0.81
LJWC + NRC hashtag sentiment + opinion lexicons + emojis + sentiment emojis + categorized emojis + colored emojis	72.57	0.72	65.86	0.63	73.76	0.73	74.93	0.74	90.04	0.90	82.87	0.82
LJWC + 3 $\geq$ gram ≥ 1 + NRC hashtag sentiment + opinion lexicons + emojis + sentiment emojis + categorized emojis + colored emojis	93.10	0.93	90.62	0.90	89.46	0.89	72.29	0.71	92.52	0.92	85.74	0.85

DT by 1.02%. The best classifier in identifying the life aspect is LR with 60.48% accuracy as Table 10 illustrated.

New Zealand Terrorist Attacks

In this case study scenario, we analyze psychological needs for New Zealand terrorist attack event to recognize public reactions based on event evolution and authorities' reactions. The Christchurch terrorist attack is a violent religion-based attack which occurred on March 15th, 2019. An Australian gunman orchestrated two consecutive mass shooting attacks on mosques during prayer in the city of Christchurch, New Zealand. 51 worshipers were killed and more than 40 injured. According to the news, the shooter published a racist manifesto detailing his motivations for the attacks on social media platforms. On Friday morning, he posted a tweet detailing his intentions to attack the two mosques saying "I will carry out an attack against the invaders and will even live stream the attack via Facebook". Later in that day, he did, indeed, live stream the first attack using Facebook's live service.

Psychological Human Need Analysis

We analyzed people's psychological needs during this event to illustrate how the dynamic evolution of the event (investigations, updates, authorities' responses and support) can affect and change a population's satisfaction levels. We collected tweets during the day of the event, as well as the following 6 days. We utilize all the possible event-related hashtags such as #ChristchurchMosqueAttack, #NewZealandTerroristAttack and #NewZealand to retrieve tweets. This step allowed us to gather in the neighborhood of 85,000 tweets along with their attached images and meta-data, including the post's geo-tagged location and time stamp. These tweets then went through the Data Preprocessing Module and Feature Extraction Module in each of the framework layers. The processed tweets were then fed to the NCR model to filter out useless tweets that did not express human needs. Around 13,675 tweets out of 85,000 were consequently ignored. The remaining 71,325 useful tweets which express people psychological needs were further analyzed to (1) identify need type using NTI model, (2) the satisfaction level using NSM model, (3) to evaluate the social context surrounding people using SCE model and (4) to identify life aspect of the expressed need using the LAI model. For a more in-depth analysis, we applied the SNIE and the FNIE models to determine the intensity score of the satisfaction level.

This analysis shows that, throughout this event, people expressed the relatedness need most, at 63.61%. The competence need followed at 19.57%, and autonomy need the

Fig. 10 NSM model accuracy before and after selecting the most predictive features using GainRatio

least at 16.81%. As Fig. 17 shows, among the 7 days of the event, tweets representing relatedness need appeared more on March 15th, 16th and 20th. Competence need-related tweets appeared more on March 18th and 19th, where autonomy need emerged more on March 17th. As Fig. 18 shows, tweets posted during this event reveal that people feel frustrated 56.73% of the time, where only 29.06% of

tweets expressed satisfied needs. The latter number rises on specific days as a reaction to certain happenings, which will be demonstrated in the detail below. Roughly 32.89% of tweets were classified to not have a clear satisfaction level and were, therefore, not considered in further analysis. Deeper analysis reflected a high frustration level of the relatedness need at 30.87%, among other need types such

Fig. 11 The effect of different textual features on the accuracy of the SNIE and FNIE models

Table 7 SNIE model performance, comparing the use of different features

Features	Pearson correlation r	RMSE
LIWC	0.58	0.17
NRC-Hash-Emo + SentiStrength	0.59	0.165
LIWC + NRC affect intensity	0.59	0.165
LIWC + NRC affect intensity + SentiStrength	0.599	0.165
LIWC + NRC affect intensity + SentiStrength + AFINN	0.60	0.164
LIWC + NRC affect intensity + SentiStrength + AFINN + Sentiment140	0.62	0.16
LIWC + NRC affect intensity + SentiStrength + AFINN + Sentiment140 + NRC-Hash-Emo	0.627	0.16
LIWC + NRC affect intensity + SentiStrength + AFINN + Sentiment140 + NRC-Hash-Emo + SentiWordNet	0.625	0.16
LIWC + NRC affect intensity + SentiStrength + AFINN + Sentiment140 + NRC-Hash-Emo + SenticNet + Senti-WordNet + Sentiment Emojis	0.641	0.158
LIWC + NRC affect intensity + SentiStrength + AFINN + Sentiment140 + NRC-Hash-Emo + SenticNet + Senti-WordNet + Sentiment Emojis + 3 > gram > 1	0.71	0.14

Table 8 FNIE model performance, comparing the use of different features

Features	Pearson correlation r	RMSE
LIWC	0.51	0.16
NRC-Hash-Emo + SentiStrength	0.43	0.16
LIWC + NRC affect intensity	0.52	0.16
LIWC + NRC affect intensity + SentiStrength	0.54	0.15
LIWC + NRC affect intensity + SentiStrength + AFINN	0.54	0.15
LIWC + NRC affect intensity + SentiStrength + AFINN + Sentiment140	0.57	0.15
LIWC + NRC affect intensity + SentiStrength + AFINN + Sentiment140 + NRC-Hash-Emo	0.58	0.15
LIWC + NRC affect intensity + SentiStrength + AFINN + Sentiment140 + NRC-Hash-Emo + SentiWordNet	0.58	0.15
LIWC + NRC affect intensity + SentiStrength + AFINN + Sentiment140 + NRC-Hash-Emo + SenticNet + SentiWordNet + sentiment emojis	0.59	0.15
LIWC + NRC affect intensity + SentiStrength + AFINN + Sentiment140 + NRC-Hash-Emo + SenticNet + SentiWordNet + sentiment emojis + 3 > gram > 1	0.68	0.13

Table 9 Pearson correlation r and RMSE for the final FNIE and SNIE models

Intensity estimator	Regression model	r	RMSE
SNIE model	SVR	0.73	0.14
FNIE model	SVR	0.72	0.13

as competence need (7.43%) and autonomy need (5.28%), as shown in Fig. 19.

Event progression including police investigation updates, authorities' responses, and surrounding support all affect human satisfaction levels; therefore, we used the same approach to create a timeline-based textual and visual representation to describe the event. We also relied on the news to understand the reasoning behind the use of some ambiguous hashtags and keywords which appeared frequently in the tweets. Figures 20 and 21 illustrate the changes in need satisfaction levels for the identified need types of autonomy, relatedness and competence throughout the event.

On March 15th, the day of the attacks, most tweets posted expressed a high frustration level for all the three needs:

relatedness 67.14%, autonomy 43.7% and competence 55%. Less than half of the tweets expressed satisfied needs. On this day, people immediately started categorizing the event as a religious/racist-related attack, as reflected in the word cloud (a) in Fig. 22. They used hashtags such as #muslims, #terrorist, #islamophobia, #breakingnews, #muslimsarenotterrorist and #brentonTarrant (shooter's name), as can be seen in Table 13. They expressed their prayers and condolences to the victims and their families using #pray-forNewZealand and #rip hashtags. Also, people used #Facebook to talk about the shared livestreamed video for the two attacks. Through the news, we came to understand the usage of #helloworldbrother, which was among the top ten hashtags. "Hello Brother" were the last words of the first victim of the shooter before he was shot and killed. An image with the "Hello Brother" title was the most shared image on that day.

Frustration levels for all the three needs escalated to their highest on the second day, March 16th, where relatedness needs reached 71%, competence 58.05% and autonomy 54.8%. The satisfaction level deteriorated on this day, especially for autonomy need (40.18% to 39.57%). Hashtags such as #whitesupremacist, #whiteprivilege, #whiteterrorism and

Fig. 12 SNIE and FNIE model accuracy before and after selecting the most predictive features using GainRatio

#whitesupremacy were among the most frequent hashtags describing the shooter's motivation behind the attack. Moreover, #49lives emerged that day to count and update the loss of lives. As Fig. 23 shows, the shooter's image was prominent in the image collections on the second day, as well as another magazine picture depicting similar attack events in different ways.

On March 17th, there was a noticeable change in the need satisfaction levels. Tweets expressing dissatisfied needs dropped significantly for all the three needs: 11.47% in relatedness needs, 15.38% in autonomy and 15.15% in competence needs, as Fig. 5 shows. Moreover, in Fig. 24, we can see that the usage of dissatisfaction expressions with high

and moderate intensity level in tweets presented less than the previous days. There was also an increase in the number of tweets expressing satisfied needs for all the three need categories, reflecting less intense emotional expressions, as Fig. 25 shows. Satisfied relatedness needs increased by 5.18%, competence need increased by 13.8% and autonomy need increased by 14.39%.

As Table 13, word cloud (c) in Fig. 22 shows, new topics of conversation arose centering around the Prime Minister's responses during her statement using #jacinda arden, #theyareus and #50lives. Moreover, #eggboy and #fraseranning trending on Twitter on that day garnered worldwide attention. People shared their feelings and opinions about the

THE EFFECT OF TEXTUAL FEATURES ON (SCE) MODEL

Fig. 13 The effect of different textual features on the accuracy of the SCE model

Fig. 14 SCE model accuracy before and after selecting the most predictive features using GainRatio

Table 10 Accuracy, recall (R), precision (P) and F_{score} for the final NCR, NTI, NSM, SCE and LAI models

Psychological need recognition models	Best classifier	Accuracy %	R	P	F
Need content recognition (NCR) model	MNB	79.13	0.791	0.82	0.77
Need type identification (NTI) model	SVM	81.97	0.81	0.82	0.817
Need satisfaction level measurement (NSM) model	SVM	93.56	0.93	0.93	0.93
Social context evaluation (SCE) model	RF	75.70	0.75	0.76	0.75
Life aspect identification (LAI) model	LR	60.48	0.59	0.60	0.59

Table 11 Accuracy and F_{score} of the SCE model, comparing the use of different features

Features	SVM		MNB		LR		K-NN		RF		DT	
	A	F	A	F	A	F	A	F	A	F	A	F
LIWC	64.72	0.63	62.61	0.62	64.61	0.64	62.88	0.60	70.50	0.70	60.04	0.59
LIWC + LCM (DAVs, IAV, SV)	64.61	0.63	62.46	0.61	64.77	0.64	63.61	0.61	71.45	0.71	60.41	0.60
LIWC + emojis	65.08	0.64	62.61	0.61	64.77	0.64	62.88	0.60	71.81	0.71	60.19	0.60
Emotion word embeddings (EWE)	65.45	0.65	48.63	0.38	64.87	0.64	59.83	0.59	65.82	0.65	49.36	0.49
NRC word-emotion + NRC hashtag emotion	62.30	0.61	56.94	0.56	61.30	0.61	60.67	0.60	65.61	0.65	59.04	0.58
(MPQA, Bing Liu, NRC) Sentiment lexicons + sentiment emojis	57.30	0.54	52.36	0.52	56.46	0.55	57.67	0.57	62.14	0.61	62.09	0.62
NRC word-emotion + NRC hashtag emotion + (MPQA, Bing Liu, NRC) sentiment lexicons	64.40	0.63	58.88	0.58	63.14	0.62	58.25	0.57	70.66	0.70	61.98	0.61
LIWC + (MPQA, Bing Liu, NRC) sentiment lexicons + sentiment emojis	65.98	0.65	61.88	0.60	65.14	0.64	62.93	0.60	71.81	0.71	61.30	0.61
LIWC + NRC word-emotion + NRC hash emotion	65.35	0.64	63.88	0.63	65.29	0.65	65.35	0.63	72.18	0.71	61.82	0.61
LIWC + NRC word-emotion + NRC hash emotion + (MPQA, Bing Liu, NRC) sentiment lexicons	65.56	0.65	63.82	0.62	64.93	0.64	64.61	0.62	71.97	0.71	60.83	0.60
LIWC + NRC word-emotion + NRC hashtag emotion + (MPQA, Bing Liu, NRC) sentiment lexicons + emoji	66.56	0.65	62.82	0.61	64.66	0.64	64.87	0.63	71.92	0.71	61.51	0.61
LIWC + 3 ≥ gram ≥ 1	70.76	0.70	71.92	0.71	68.19	0.67	59.83	0.57	74.65	0.74	65.72	0.65
LIWC + LCM (DAVs, IAV, SV) + 3 ≥ gram ≥ 1	70.60	0.69	71.97	0.71	68.45	0.68	59.72	0.57	75.70	0.75	64.87	0.64
LIWC + 3 ≥ gram ≥ 1 + NRC word-emotion + NRC hashtag emotion	70.66	0.70	72.02	0.71	69.13	0.68	59.41	0.56	74.55	0.74	63.82	0.63
LIWC + 3 ≥ gram ≥ 1 + NRC word-emotion + NRC hashtag emotion + (MPQA, Bing Liu, NRC) sentiment lexicons	71.03	0.70	71.50	0.70	68.29	0.67	59.51	0.56	74.76	0.74	63.88	0.63
LIWC + 3 ≥ gram ≥ 1 + NRC word-emotion + NRC hashtag emotion + (MPQA, Bing Liu, NRC) sentiment lexicons + emoji + sentiment emojis	70.87	0.70	71.60	0.70	69.45	0.69	59.51	0.56	73.86	0.73	65.03	0.65
LIWC + 3 ≥ gram ≥ 1 + NRC word-emotion + NRC hashtag emotion + (MPQA, Bing Liu, NRC) sentiment lexicons + emoji + sentiment emojis + POS	71.13	0.70	72.87	0.71	69.24	0.68	59.51	0.56	73.23	0.73	64.56	0.64
LIWC + LCM (DAVs, IAV, SV) + 3 ≥ gram ≥ 1 + NRC word-emotion + NRC hashtag emotion + negation lexicon + (MPQA, Bing Liu, NRC) sentiment lexicons + emoji + sentiment emojis + POS + emotion word embeddings (EWE)	70.87	0.70	72.34	0.72	69.50	0.69	60.98	0.58	68.76	0.68	64.03	0.64

Fig. 15 The effect of different textual features on the accuracy of the LAI model

Table 12 Accuracy and F_{score} of the LAI model, comparing the use of different features

Features	SVM		MNB		LR		K-NN		RF		DT	
	A	F	A	F	A	F	A	F	A	F	A	F
LIWC	51.69	0.48	39.06	0.33	51.63	0.49	36.97	0.35	51.80	0.48	39.79	0.39
LIWC + life aspect lexicons	54.05	0.51	41.37	0.36	53.04	0.51	37.31	0.35	54.73	0.51	42.33	0.42
3 ≥ gram ≥ 1 + LIWC	54.73	0.54	55.97	0.54	56.82	0.56	42.16	0.39	54.17	0.50	48.13	0.47
3 ≥ gram ≥ 1 + life aspect lexicons	55.07	0.54	57.72	0.56	54.67	0.53	45.26	0.42	55.46	0.53	46.22	0.45
3 ≥ gram ≥ 1 + LIWC + life aspect lexicons	55.52	0.55	56.70	0.55	56.03	0.55	43.12	0.39	55.18	0.51	49.43	0.48

Fig. 16 LAI model accuracy before and after selecting the most predictive features using GainRatio

Fig. 17 The need types identified during the New Zealand terrorist attacks using the NTI model

Fig. 18 Measuring satisfaction levels using the NSM model during the days of the New Zealand terrorist attacks

Fig. 19 Identifying need types present during the New Zealand terrorist attacks using the NTI model, and measuring satisfaction levels using the NSM model

Fig. 20 Measuring the changes in need frustration levels during the day of the New Zealand terrorist attacks and the subsequent 6 days

Fig. 21 Measuring the changes in need satisfaction levels during the day of the Zealand terrorist attacks and the subsequent 6 days

Fig. 23 A chronological visual representation for the New Zealand terrorist attacks

Fig. 24 Intensity levels of frustrated needs measured during the New Zealand terrorist attacks using the FNIE model

Fig. 25 Intensity levels of satisfied needs measured during the New Zealand terrorist attacks using the SNIE model

Fig. 26 Word cloud generated from tweets posted on March, 21st

by 5.14 points, and competence needs fell by 4.1, while the dissatisfaction increased, especially for relatedness needs, by 5.13%. On that day, as can be seen from Table 13, #guns, #hijab, #Utrecht and #guncontrol are new hashtags, leading the subtopic of discussion. On the fifth day after the shooting, New Zealand Prime Minister, Jacinda Ardern, announced a plan to reform gun law, banning all military-style semi-automatic weapons. This may explain the higher expressions of satisfied autonomy needs and the lower expression of dissatisfied autonomy need tweets posted on this day. Moreover, #hijab was used more frequently, with non-Muslim women showing their care and support. #Utrecht was used frequently on March 19th; furthermore, it climbed to be one of the ten most frequently used hashtags on this day. People used #Utrecht in reference to another shooting attack in Utrecht, Netherlands, which occurred on March 18th, 2019, where a man opened fire on a tram killing three people and injuring seven. Tweets posted on this day mentioned the two attacks, comparing how the news and media reported them differently based on the religion and nationality of the attackers. The visual representation in Fig. 23 shows how images reflecting support and solidarity continue to lead the visual representation.

On Thursday, March 21st, within all three need categories, people satisfaction levels increased, and their frustration levels decreased. They also expressed less moderate and high-intensity frustration words and more

low intensive language compared to the beginning of the event, as Fig. 24 shows. New sub-topics mentioned that day were #respect, #peace, #headscarfforharmony, #leadership and #banassaultweapons (Fig. 26). On this day, women announced a “scarves in solidarity” event to be hold on Friday using the hashtags #headscarfforharmony, #respect, #peace and #leadership hashtags were mentioned frequently in tweets talking about the NZ prime minster’s speech and how people in New Zealand supported and dealt with that tragic event.

Satisfaction Level Intensity Estimation

Gauging the intensity of the language used in posts to express satisfaction and frustration needs gave us a deeper understanding of the population’s feelings and reactions in regard to their levels of satisfaction versus frustration. As we can see from Fig. 24, people used high and moderate emotional expressions to convey their frustration level, which revealed how angry they were throughout the event. On the other hand, they used expressions between low- and moderate-intensity levels to communicate their satisfaction level. Very few high-intensity satisfaction expressions were reflected during this event, as Fig. 25 shows.

Social Context Analysis

For social context analysis, Fig. 27 shows how people in New Zealand observe their surroundings and evaluate their social context. We notice that the day of the event and the next day, people reported that they did not perceive a supportive environment, at percentages of 30.2% on March 15th, and 31.7% on March 16th, which are the highest when compared to the other days. This confirms that their frustration was at its highest level in the first 2 days, as shown in Fig. 18. Tweets reflecting that people identify their environment as non-supportive appeared less often by March 17th and March 20th. The highest volume of tweets expressing supportive social context came on March 20th. On that day, and for the first time, people posted tweets indicating

Fig. 27 Using the SCE model to identify social context types during the New Zealand terrorist attacks

Table 14 The top ten most frequently used hashtags classified under supportive social context and non-supportive social context posted on March 15th, March 17th and March 21st during the New Zealand terrorist attacks event

March 15th		March 17th		March 21st	
Supportive social context	Non-supportive social context	Supportive social context	Non-supportive social context	Supportive social context	Non-supportive social context
#Auspol	#Immigrant	#EggBoy	#Auspol	#Gun	#BackStronger
#ABC	#Muslims	#Muslims	#50lives	#JacindaArdern	#Boycott
#Kiwis	#Australian	#EggBoyForPM	#FarRightExtremism	#KiaKaha	#FraserAnning
#NewZealand	#MosqueMassacre	#JacindaArdern	#Islamophobia	#NZFirst	#Hate
#PennyWong	#PeacefulMosques	#NaeemRashid	#Muslims	#StoptheRacism	#HeartBreaking
#WeLoveOurPeople	#Terrorism	#TheyAreUs	#NZMosqueShooting	#TheyAreUs	#Politics
#WeStandTogether	#Auspol	#Kindness	#Funerals	#StandTogethe	#Terrorist
#Respect	#Chchshooting	#Love	#Trump	#FightHateWithLove	#TheyAreUs
#TheMuslimPhotographer	#Respect	#Solidarity	#NZpol	#2MinutesSilence	#NZGovernment
#WeAreYou	#TheyAreUs	#StandTogether	#White	#Auckland	#Banned

Fig. 28 Word cloud generated from tweets which were posted on March 15th (a, b), March 17th (c, d) and March 21st (e, f), during the New Zealand terrorist attack, and were identified to have supportive social context (a, c, e) and non-supportive social context (b, d, f)

that they felt they had a supportive environment more than non-supportive environment. This affirms our statement that certain sub-events during these days, such as the Prime Minister's official speech on March 17th, and the announcement of a plan for gun law reform on March 20th, could be the reasons for the elevation in tweets indicating supportive

social context. Table 14 shows the most frequent sub-topics (hashtags) posted on March 15th, 17th and 20th. The word cloud in Fig. 28 includes the most frequent words used on these days (a) and (b) for March 15th, (c) and (d) for March 17th and (e) and (f) for March 21st.

Fig. 29 Identified life aspects extracted from tweets posted during New Zealand terrorist attacks event

Location-Based Need Analysis

As a situation unfolds, people analyze event details based on their points of view, which are influenced by different factors such as age, gender, race, culture, religion, interests and many others. These factors have a significant impact on the way people perceive any particular situation and how it affects them. Accordingly, the posts gathered during the New Zealand incident bring to light the diverse angles from which each individual's perception is generated. This diversity is illustrated through the tweet categorization in Fig. 29. An individual's geographical location and cultural observations may have an impact on needs. Therefore, we analyze need satisfaction level based on location (where they were when they posted the tweets sharing their feelings and needs). Figure 30a, b shows location-based analysis for need satisfaction levels. We selected Canada and the United States

Fig. 30 Measuring individuals' satisfaction levels using tweets generated from within Canada, United States and Middle East during the New Zealand terrorist attacks

Table 15 The top ten most frequently used hashtags identified as reflecting satisfied needs or dissatisfied needs posted within Canada, United States and Middle East

Middle East		Canada and United states	
Sub-topics for satisfied needs	Sub-topics for frustrated needs	Sub-topics for satisfied needs	Sub-topics for frustrated needs
#HelloBrother	#HelloBrother	#Muslims	#WhiteSupremacist
#Muslims	#Terrorist	#EggBoy	#WhitePrivilege
#PrayforChristchurch	#Muslims	#WhiteSupremacist	#Islamophobia
#يهاهوالا - دناليزوين - موج	#Islamophobia	#prayforchristchurch	#WhiteSupremacy
#Islamophobia	#Media	#Islamophobia	#BreakingNewsNow
#Istanbul	#WhiteSupremacist	#AllahuAkbar	#WhiteTerrorism
#JacindaArdern	#يهاهوالا - دناليزوين - ثداح	#CAIR	#Muslim
#Jesus	#Australia	#PrayForChristchurch	#Terrorist
#Kuwait	#BrentonTarrant	#OneTeamOneFight	#FacebookMassacre
#Quran	#Islam	#JacindaArdern	#WhiteNationalist

(combined) and Middle East countries. In Canada and the United States, 32.39% of tweets reveal that people were satisfied, whereas 52.23% of tweets express frustrated needs. In Middle East countries, 28.01% of the posted content indicates that people felt satisfied, where 55.95% shows frustrated needs. Table 15 shows the most frequent sub-topics discussed when population needs were satisfied vs frustrated in the two locations during the first 3 days of the event.

Conclusion and Future Work

The interpretation of data garnered through the proposed psychological need recognition and analysis can assist authorities providing a heightened situational awareness and improving management of pre- and post-event conflicts and social reactions. These methods have been proven to enrich the effectiveness of smart city research and development. In this work, we design, implement, and evaluate a theoretical-based multi-layered reference model, which is capable of identifying human psychological needs, and ascertaining their satisfaction level, evaluating people surrounding environment with regard to various life aspects. The design and the development of the reference model layers are influenced by motivational psychology research. Various linguistic, psychological and Twitter-based features are explored in conjunction with distinctive machine learning algorithms to develop psychological need models based on the conceptual layered reference model. These include (1) the need content recognition (NCR) model, which is used to recognize need content, (2) the need type identification (NTI) model, to identify the need type, (3) the need satisfaction level measurement (NSM) model, which measures individual need satisfaction level or whether the detected need is satisfied or frustrated, (4) the social context evaluation (SCE) model to evaluate the individual surrounding environment and whether they are supportive or non-supportive, and, (5) the life aspect identification (LAI) model which is employed to identify the life domain. Furthermore, the frustrated need intensity estimator (FNIE) model and satisfied need intensity estimator (SNIE) model are implemented within the third layer to obtain the intensity score of the satisfaction level. The results of our experiments affirm the effectiveness of the proposed framework and the developed psychological needs models. We implemented a prototype of the proposed framework in a real-life case study setting, analyzing population needs during the New Zealand terror attacks, which occurred on March 15th, 2019. The analysis of these critical events gave a clear indication of the feasibility of applying the proposed framework and its effectiveness in detecting changes among public reaction in terms of their psychological needs based on event

evolution and authority response. The proposed framework could potentially be employed in a variety other application such as marketing and need-based recommendation scenarios. For future work, we aim to enrich the framework with other types of psychological needs. Extending the capability of the framework by integrating multilingual aspects for need analysis in other commonly spoken languages to cover a wide range of geographical locations is another possible future expansion.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Pellicer S, Santa G, Bleda AL, Maestre R, Jara AJ, Skarmeta AG. A global perspective of smart cities: a survey. In: Seventh international conference on innovative mobile and internet services in ubiquitous computing. IEEE; 2013. p. 439–44.
2. Hall RE, Bowerman B, Braverman J, Taylor J, Todosow H, Von Wimmersperg U. The vision of a smart city. Upton: Brookhaven National Lab. Technical report; 2000.
3. Tumasjan A, Sprenger TO, Sandner PG, Welpe IM. Predicting elections with twitter: what 140 characters reveal about political sentiment. In: Proceedings of the fourth international AAAI conference on weblogs and social media (ICWSM), vol. 10. 2010. p. 178–85.
4. Coe A, Paquet G, Roy J. E-governance and smart communities: a social learning challenge. *Soc Sci Comput Rev*. 2001;19(1):80–93.
5. Guthier B, Abaalkhail R, Alharthi R, El Saddik A. The affect-aware city. In: International conference on computing, networking and communications (ICNC). IEEE; 2015. p. 630–36.
6. Pak A, Paroubek P. Twitter as a corpus for sentiment analysis and opinion mining. In: International conference on language resources and evaluation (LREc), vol. 10. 2010. p. 1320–6.
7. Kouloumpis E, Wilson T, Moore JD. Twitter sentiment analysis: the good the bad and the omg! In: Fifth international AAAI conference on weblogs and social media (ICWSM), vol. 11. 2011.
8. Balabantaray RC, Mohammad M, Sharma N. Multi-class twitter emotion classification: a new approach. *Int J Appl Inf Syst*. 2012;4(1):48–53.
9. Yan JLS, Turtle HR. Exploring fine-grained emotion detection in tweets. In: Proceedings of the North American chapter of the

- Association for Computational Linguistics: human language technologies (NAACL-HLT). 2016. p. 73–80.
10. Golbeck J, Robles C, Turner K. Predicting personality with social media. CHI'11 extended abstracts on human factors in computing systems. New York: ACM; 2011. p. 253–62.
 11. Bollen J, Pepe A, Mao H. Modeling public mood and emotion: twitter sentiment and socio-economic phenomena. In: Proceedings of the fifth international AAAI conference on weblogs and social media (ICWSM), vol. 11. 2011. p. 450–453.
 12. Milyavskaya M, Koestner R. Psychological needs, motivation, and well-being: a test of self-determination theory across multiple domains. *Personal Individ Differ*. 2011;50(3):387–91.
 13. Sheldon KM, Schuler J. Wanting, having, and needing: integrating motive disposition theory and self-determination theory. *J Personal Soc Psychol*. 2011;101(5):1106.
 14. Tay L, Diener E. Needs and subjective well-being around the world. *J Personal Soc Psychol*. 2011;101(2):354.
 15. La Guardia J, Ryan R, Couchman C, Deci E. Basic psychological needs scales. *J Personal Soc Psychol*. 2000;79:367–84.
 16. Van den Broeck A, Vansteenkiste M, De Witte H, Soenens B, Lens W. Capturing autonomy, competence, and relatedness at work: construction and initial validation of the work-related basic need satisfaction scale. *J Occup Organ Psychol*. 2010;83(4):981–1002.
 17. Klassen RM, Perry NE, Frenzel AC. Teachers' relatedness with students: an underemphasized component of teachers' basic psychological needs. *J Educ Psychol*. 2012;104(1):150.
 18. Yang H, Li Y. Identifying user needs from social media. San Jose: IBM Research Division; 2013. p. 11.
 19. Max-Neef M, Elizalde A, Hopenhayn M. Development and human needs. In: *Real-life economics: understanding wealth creation*. London, New York: Routledge; 1992. p. 197–213.
 20. Wahba MA, Bridwell LG. Maslow reconsidered: a review of research on the need hierarchy theory. *Organ Behav Hum Perform*. 1976;15(2):212–40.
 21. Kenrick DT, Griskevicius V, Neuberg SL, Schaller M. Renovating the pyramid of needs: contemporary extensions built upon ancient foundations. *Perspect Psychol Sci*. 2010;5(3):292–314.
 22. Rashkin H, Bosselut A, Sap M, Knight K, Choi Y. Modeling naive psychology of characters in simple commonsense stories. 2018. arXiv preprint. [arXiv:1805.06533](https://arxiv.org/abs/1805.06533).
 23. Ding H, Jiang T, Riloff E. Why is an event affective? Classifying affective events based on human needs. In: Proceedings of Association for the Advancement of Artificial Intelligence (AAAI) workshop on affective content analysis. 2018. p. 8–15.
 24. Deci EL, Ryan RM. The “what” and “why” of goal pursuits: human needs and the self-determination of behavior. *Psychol Inq*. 2000;11(4):37–41.
 25. Danielsen G. Meeting human needs, preventing violence: applying human needs theory to the conflict in sri lanka. Unpublished MA thesis, Universidad del Salvador, Buenos Aires; 2005.
 26. Deci EL, Ryan RM. Self-determination theory: a macrotheory of human motivation, development, and health. *Can Psychol*. 2008;49(3):182.
 27. Chen B, Vansteenkiste M, Beyers W, Boone L, Deci EL, Van der Kaap-Deeder J, Duriez B, Lens W, Matos L, Mouratidis A, et al. Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motiv Emot*. 2015;39(2):216–36.
 28. Sheldon KM, Elliot AJ, Kim Y, Kasser T. What is satisfying about satisfying events? Testing 10 candidate psychological needs. *J Personal Soc Psychol*. 2001;80(2):325.
 29. Britton PC, Van Orden KA, Hirsch JK, Williams GC. Basic psychological needs, suicidal ideation, and risk for suicidal behavior in young adults. *Suicide Life Threat Behav*. 2014;44(4):362–71.
 30. Deci EL, Ryan RM. Motivation, personality, and development within embedded social contexts: an overview of self-determination theory. In: *The Oxford handbook of human motivation*. Oxford University Press; 2012. p. 85–107.
 31. Gagné M. The role of autonomy support and autonomy orientation in prosocial behavior engagement. *Motiv Emot*. 2003;27(3):199–223.
 32. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. *J Lang Soc Psychol*. 2010;29(1):24–54.
 33. Semin GR, Fiedler K. The linguistic category model, its bases, applications and range. *Eur Rev Soc Psychol*. 1991;2(1):1–30.
 34. Coenen L, Hedeboom L, Semin GR. The linguistic category model (lcm). Manual. Amsterdam: Free University Amsterdam; 2006. p. 1 151 434 261 594–8567.
 35. Pavalanathan U, Eisenstein J. Emoticons vs. emojis on twitter: a causal inference approach. 2015. arXiv preprint. [arXiv:1510.08480](https://arxiv.org/abs/1510.08480).
 36. Novak PK, Smailović J, Sluban B, Mozetič I. Sentiment of emojis. *PLoS One*. 2015;10(12):e0144296.
 37. Baran G. Nonviolent communication: an important component in personal and nonviolent social change. *Acorn*. 2000;10(2):42–8.
 38. Hu M, Liu B. Mining and summarizing customer reviews. In: *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining*. ACM; 2004. p. 168–77.
 39. Potts C. On the negativity of negation. In: *Proceedings of semantics and linguistic theory*, vol. 20. 2010. p. 636–59.
 40. Svetlana Kiritchenko XZ, Mohammad SM. Sentiment analysis of short informal texts. *J Artif Intell Res*. 2014;50:723–62.
 41. Wilson T, Wiebe J, Hoffmann P. Recognizing contextual polarity in phrase-level sentiment analysis. In: *Proceedings of human language technology conference and conference on empirical methods in natural language processing*. 2005. p. 347–54.
 42. Mohammad SM, Turney PD. Crowdsourcing a word-emotion association lexicon. *Comput Intell*. 2013;29(3):436–65.
 43. Baccianella S, Esuli A, Sebastiani F. SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: *International conference on language resources and evaluation (LREC)*, vol. 10. 2010. p. 2200–4.
 44. Nielsen FÅ. A new anew: evaluation of a word list for sentiment analysis in microblogs. In: *Proceedings of the ESWC2011 workshop on 'Making Sense of Microposts': big things come in small packages*. 2011. arXiv preprint. [arXiv:1103.2903](https://arxiv.org/abs/1103.2903).
 45. Bradley MM, Lang PJ. Affective norms for english words (anew): instruction manual and affective ratings. Technical report C-1. The Center for Research in Psychophysiology, University of Florida; 1999.
 46. Mohammad S, Kiritchenko S, Zhu X. NRC-canada: building the state-of-the-art in sentiment analysis of tweets. In: *Proceedings of the seventh international workshop on semantic evaluation exercises (SemEval)*. 2013. arXiv preprint. [arXiv:1308.6242](https://arxiv.org/abs/1308.6242).
 47. Mohammad SM. Word affect intensities. In: *Proceedings of the 11th edition of the language resources and evaluation conference (LREC)*. 2018. arXiv preprint. [arXiv:1704.08798](https://arxiv.org/abs/1704.08798).
 48. Cambria E, Poria S, Hazarika D, Kwok K. Senticnet 5: discovering conceptual primitives for sentiment analysis by means of context embeddings. In: *Thirty-second AAAI conference on artificial intelligence (AAAI)*. 2018.
 49. Gimpel K, Schneider N, O'Connor B, Das D, Mills D, Eisenstein J, Heilman M, Yogatama D, Flanigan J, Smith NA. Part-of-speech tagging for twitter: annotation, features, and experiments. Carnegie-Mellon Univ Pittsburgh Pa School of Computer Science, Technical Report; 2010.

50. Qi Y, Sachan DS, Felix M, Padmanabhan SJ, Neubig G. When and why are pre-trained word embeddings useful for neural machine translation? 2018. arXiv preprint. [arXiv:1804.06323](https://arxiv.org/abs/1804.06323).
51. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. 2013. arXiv preprint. [arXiv:1301.3781](https://arxiv.org/abs/1301.3781).
52. Agrawal A, An A, Papagelis M. Learning emotion-enriched word representations. In: Proceedings of the 27th international conference on computational linguistics. 2018. p. 950–61.
53. Pennington J, Socher R, Manning C. Glove: global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). 2014. p. 1532–43.
54. Leyba T. Semantic search by means of word sense disambiguation using a lexicon. US Patent 9,317,589; 2016.
55. Alharthi R, Guthier B, El Saddik A. Recognizing human needs during critical events using machine learning powered psychology-based framework. *IEEE Access*. 2018;6:58737–53.
56. Alharthi R, Guthier B, Guertin C, El Saddik A. A dataset for psychological human needs detection from social networks. *IEEE Access*. 2017;5:9109–17.
57. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. Smote: synthetic minority over-sampling technique. *J Artif Intell Res*. 2002;16:321–57.
58. Mohammad SM, Bravo-Marquez F. Emotion intensities in tweets. In: Proceedings of the sixth joint conference on lexical and computational semantics (*Sem). Vancouver; 2017. arXiv preprint. [arXiv:1708.03696](https://arxiv.org/abs/1708.03696).
59. Karegowda AG, Manjunath A, Jayaram M. Comparative study of attribute selection using gain ratio and correlation based feature selection. *Int J Inf Technol Knowl Manag*. 2010;2(2):271–7.
60. Zinkevich M. Rules of machine learning: best practices for ml engineering. Google. [Online]. 2017. <https://developers.google.com/machine-learning/guides/rules-of-ml/>. Accessed 18 Oct 2018.
61. Hinton G, Deng L, Yu D, Dahl GE, Mohamed A-R, Jaitly N, Senior A, Vanhoucke V, Nguyen P, Sainath TN, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Process Mag*. 2012;29(6):82–97.
62. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. 2012. p. 1097–105.
63. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The weka data mining software: an update. *ACM SIGKDD Explor Newsl*. 2009;11(1):10–8.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.