Emotion Analysis Using Machine Learning Classifiers in Tourism and Management: A Comparative Study Utilizing Kaggle Dataset and MATLAB

Harshita Mishra[1]*, Ashwani Kumar[2], Parveen Kumar Saini[3]

[1, 2&3] Department of Computer Science and Engineering, Chandigarh University, Mohali, India

* Corresponding author- mishraharshita0412@ gmail.com

Abstract. A crucial component in comprehending human – computer interaction is emotion analysis, which helps create systems that are more responsive and intuitive. In this works, we investigate several classifiers for efficient textual data- based emotion identification. Several classifiers are examined in the study such as decision trees, linear quadratic, Support vector machines, Confusion Matrix, and SVMs. Accurate classification of emotional states by each classifier was assessed based on its performance. Our results demonstrate notable variations in the classifiers accuracy and computational efficiency, with certain SVM configurations performing better than others. The studies findings emphasize the potential of advanced machine learning approaches to improve human-computer interaction and further the field's continuous achievements in emotion monitoring. The knowledge acquired may be very helpful for future studies and real-world applications in variety of fields, including social media analytics, healthcare, and customer service.

Keywords – Emotions, Machine Learning, MATLAB, Classifiers.

1. Introduction

Image processing may enhance photos and extract useful data. A photograph may be effectively converted into a digital format for additional processing using this procedure. This method use a 2D image as input, with pixel values ranging from 0 to 255, much like signal processing. It involves converting an image into a two-dimensional matrix. The process entails transforming a picture to a 2D matrix (Zang et al., 2020; Sharma et al., 2021).

There are basic phases to it :-

Examining the image: An image scanned by a device produces a raw image that has to be processed. Pixels can be used to represent it, as it was previously said.

The goal of this stage is to extract computer-friendly data.

Processing and Enhancement: By sampling and quantizing the input signals, a digitizer converts the image into digital format. In order to achieve optimal quantization and resolution for the human sense of different hues using different grayscales, a high sample rate is required (Sharma et al., 2021)

The output that is produced further classifies the picture and describes its characteristics.

https://doi.org/10.52202/081568-0025

Conversion of a Color Image into a Grayscale:

There are basically two techniques to convert a color image to a grayscale one.

Average Method :- The three hues that make up a color are Red, Blue, and Green. Picture are averaged in the average technique.

Hence, Grayscale= (R+G+B)/3 is obtained. Nevertheless, sometimes the grayscale image is replaced with a black picture. This is because the altered image contains 33% of the original Red, Blue, and Green (Bandyopadhayay et al., 2012)

So, to answer this riddle, we use the second strategy, which is also referred to as the Weighted Method or Luminosity Method.

Weighted or Luminosity Method :- To solve the Average approach problem, we use the Luminosity technique. The blue hue has a proportion between these two colors, and these techniques boost the green color and diminish the presence of the red color (Guo, 2022; en et al., 2014)

Consequently, using the equation:- Grayscale=((0.3 * R) + (0.59 * G) + (0.11 * B)). The wavelength patterns of these hues are why we employ this.

Red has the greatest wavelength, and blue has the smallest.

1. **Evaluation of the Literature**

OpenCV is an abbreviation for Open Computer Vision Library. Everyone has free access to this extensive library, which contains more than 2500 algorithms designed specifically to tackle computer vision and machine learning applications (Maruf et al., 2022). These algorithms may be used to perform a wide range of tasks, such as Face Recognition, Object Identification, Camera Movement Tracking, and Scenery Recognition. It's a large community, with at least 47,000 people who actively contribute to the library (Houssein et al., 2020). Numerous public and private businesses utilize it, GPU Acceleration is one of the new features added to the current libraries. This new capability, albeit not yet completely developed, can handle most jobs. The GPU may benefit from several libraries, such as NPP, or NVIDIA performance primitives, because it is powered by CUDA. (Bota et al., 2019; Yang et al., 2020).

The advantage of GPU technology is that it can be utilized by anybody without a deep grasp of GPU programming. Rather than being able to instantly change a picture's properties in the GPU Module, we must first copy the original image and then make modifications to it (Alhalaseh & Alasasfeh, 2020)

IV. The steps for installing Python 2.7 and the required packages are as follows:

Let's begin by using an example image in the.jpg or.png format, and then use an image processing approach to extract the mood of the topic. (Every living being from which sentiments may be inferred is referred to as the "Subject").

1.1 Bringing Libraries in

To properly implement this project, you will need to download and install the NumPy, Glob, Random, and Python 2.7.x packages. The C disk is where Python installation by default takes place. Open the Python IDLE application, import each package, and begin working.

1.2 Numpy

One Python library used for complicated technical assessment is called NumPy.

https://doi.org/10.52202/081568-0025

Multidimensional arrays, which need the processing of several mathematical formulae, are implemented using it (Sharma et al., 2021)

The dimension of an array stated in a program is known as its axis.

Rank is the quantity of axes that are present in an array. For instance A= [1,2,3,4,5]

1.3 Glob

The Globo module recognizes the pattern and creates a file in relation to it based on the Unix Shell rules. The entire route name is generated.

1.3.1 Wildcards

Several actions may be carried out on files or a portion of a directory using these wildcards. There exist several functional wildcards , but only two of them are very helpful: -

Various files created in dictionaries

a) Asterisk (*): denotes any combination of any number of characters.

For instance.

import glob from glob.glob('direc/file*') for name print name; output

path/filename1 path/filename2 path/filename3 path/filename4 path/filename5 path/files

a) Question Mark (?): This indicates or points to the one missing character. The import of glob for name occurs in glob.glob('direc/filename?').

print name; output path / filename1 / filename2 / filename3/ filename4 / filename5

This wildcard is restricted to usage in a single directory; it does not expand. that is, a file in a subdirectory cannot be found (Ain et al., 2017)

D. **Indeterminate Random**

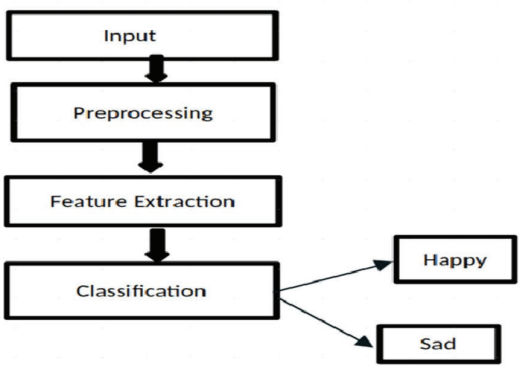
The module chooses an element, or random number, from a predefined collection of things. Certain features that provide access to these sorts of activities are compatible with this module.

E. **Indeterminate Random**

The module chooses an element, or random number, from a predefined collection of things. Certain features that provide access to these sorts of activities are compatible with this module.

Categorization of Random Module: -

First, randint(m,n) yields a value for x such that m<=x<=n Second, randrange(dog, cat, mouse, lion) Returns any element or random variable from the specified range (Doma & Pirouz, 2020)



II. V. Various emotions that can be discerned from an image include:

A. Indifferent
B. Joyful
C. Furious
D. Dismay
E. Astonishment
F. Disgust
G. Sadness
H. Fear

III. VI. Procedures for utilizing OpenCV-Python for emotion detection:

2. The first thing we need to do is create a dataset after installing all required tools. In order to ensure accuracy in our results and adequate data for information extraction, we may generate our own dataset by evaluating a series of photographs. An already-existing database can also be used.
3. Following that, the dataset is arranged into two distinct folders. Every picture will be in the first directory, and every piece of information on the various emotions will be in the second directory.
4. All of the output photographs will be saved in a different directory and arranged according to emotions before being encoded, after the execution of the Python code on the example images.
5. While OpenCV has a variety of classes for emotion identification, we will mostly employ Fisher Face one.
6. Extracting Faces: OpenCV comes with four pre-built classifiers. We utilize these classifiers sequentially to extract as many faces as we can.
7. Two subsets of the dataset are extracted: a training set and a classification set.
8. The classifier performance is estimated using the classification set, and the kind of emotion is taught using the training set, which gathers data from several photographs.
9. For optimal effects, all of the photographs should have the exact identical attributes, including size.
10. After being converted to grayscale, the topic of each photograph is scrutinized, cropped, and saved to a directory.
11. Lastly, we generate a training set using 80% of the test data and classify on the classification set using the remaining 20%.
12. Repeat the process to boost efficiency (Casale et al., 2014).

VII. Cohn-Kanade Database for AU-CODED Expressions:

This is one database that may be used to identify different moods. It is a database containing a wide variety of emotions. This database is currently divided into two categories, and a third is in the process of being created ((Doma & Pirouz, 2020).

Although the second edition is called "CK+," the first edition is called "CK." The neutral sensation is detected by the CK version before it recognizes the other higher emotions. With the introduction of this database's second kind, the frequency of processes increased by around 25% and the frequency of subjects increased by approximately 30%.

https://doi.org/10.52202/081568-0025

The third edition, which will be launched soon, will have all of the functionality found in the CK and CK+ Databases, including integrated 30 degrees of rotation from the front (Bandyopadhayay et al., 2012).

2. UTILITIES AND FUTURE RANGE

The field of computer vision is still in its early stages of development. The pace of research in this subject is accelerating (Yan, 2022).
Computer vision is not complete without the ability to identify emotions (Bandyopadhayay et al., 2012). If one can understand the complexities and final steps of a work, they can complete a lot of activities and procedures.
Several popular and extensively utilized uses of emotion detection include :

2.1 Product and app development

Testing a product's usability is one of the many software engineering procedures where emotion detection may be quite helpful in streamlining the process. It has long been known that a person's degree of comfort with various software programs greatly relies on their emotional state. Human emotions may be affected by a product's overall appearance and feel, which might influence whether or not a consumer chooses to purchase it. For those in connected industries, therefore, learning more about the many emotional states of the human body and how using certain goods affects them is crucial.

2.2 Better methods of instruction

Research indicates that although some emotional states support improved learning strategies, others attempt to repress them. It is uncommon to notice a distinction between the two groups' disparate emotional states. Positive emotional states, for instance, are considered detrimental to learning (Machova et al., 2023). whereas somewhat negative emotional states encourage critical thinking and are suitable for doing important activities.

2.3 Self-developed websites

Service providers are interested in gathering enormous amounts of data that may be taken from consumers due to the internet's massive expansion. Consequently, all of the information and adverts play according to the user's profile. Consequently, providing detailed information about the various human emotions can lead to the creation of far more accurate behavioral models representing various user types learning (Machova et al., 2023).

2.4 Video games that are fully immersive

The entertainment business includes a sizable portion of the video gaming sector. Thus, video game developers center their study on various human emotions that are frequently encountered in order to create these games much more intricate and invasive. Video games are designed with the intention of attracting an increasing number of players by seamlessly integrating human emotions into the gameplay.

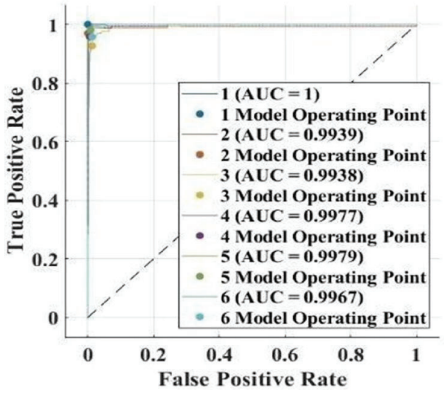
Example: Using Emotion Analysis in Hotel Review Management:
Scenario: Thousands of reviews are submitted by online users to the five-star hotel chain "Luxury Stay Hotels" via TripAdvisor and Google Reviews. In order to improve guest experiences and service quality, the hotel administration plans to examine consumer sentiments.
Data Collection: The management obtains information from a Kaggle dataset that includes text reviews, rating assessments, and lodging reviews.
Machine Learning Approach: All reviews are subjected to token separation, punctuation stripping, and stopword removal prior to analysis.

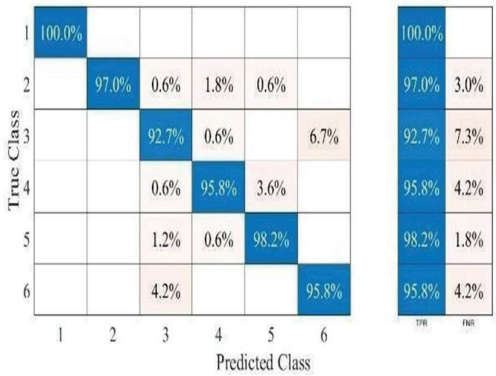
TF-IDF (Term Frequency-Inverse Document Frequency) feature extraction transforms text elements into numerical data.
Models of Classification: Support Vector Machines are used by the system to identify emotive opinions in reviews that fall into one of three categories: Favorable, neutral, or negative, random Forest for further in-depth analysis of complaints about services (Zang et al., 2020)
Textual emotion probabilities are computed in the analysis process with the aid of a Naive Bayes method.
Findings and Perspectives: Guests gave a favorable review, stating that the staff was welcoming and the rooms were immaculate, and the setting offered beautiful views of the surrounding countryside.
A modest percentage of visitors (20%) complained about the standard level of food and service.
The delayed check-in process, along with issues with Wi-Fi access and the noisy surroundings, caused 15% of guests to voice unfavorable views.

-p+
Finding different emotions in a particular sample image is the main goal. The hardest aspect of this job was figuring out which emotion was which when two of them were quite similar learning (Machova et al., 2023), For example, "disgust" may be mistaken for "sadness," "surprise" for "happy," and so on. Now, across eight distinct categories, the accuracy was almost 70%, which is really good considering how much our classifier has learnt. Thus, we need to investigate ways to improve its accuracy and efficiency.
We can see from our list of emotions that we don't have many examples of the words "sad," "fear," and "contempt." We can definitely improve performance by increasing the number of photos for these emotions. Alternatively, if we remove these feelings from the list, optimization can be improved by more than 80%.
The Cohn-Kanade (CK and CK+) Database was the dataset we used for this work.

3. CONCLUSION

This study highlights the effectiveness of various classifiers in textual emotion analysis, with SVMs demonstrating superior performance in both accuracy and computational efficiency. The findings underscore the importance of selecting appropriate machine learning models to enhance human-computer interaction. These insights contribute valuable guidance for future research and practical applications in domains such as social media monitoring, healthcare, and customer engagement. Ultimately, emotion-aware systems can lead to more intuitive and empathetic digital interactions.





https://doi.org/10.52202/081568-0025

REFERENCE

1. Zhang, J., Yin, Z., Chen, P., & Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. Information fusion, 59, 103-126.
2. Sharma, T., Diwakar, M., Singh, P., Lamba, S., Kumar, P., & Joshi, K. (2021, November). Emotion Analysis for predicting the emotion labels using Machine Learning approaches. In 2021 IEEE 8th Uttar Pradesh section international conference on electrical, electronics and computer engineering (UPCON) (pp. 1-6). IEEE.
3. Al Maruf, A., Ziyad, Z. M., Haque, M. M., & Khanam, F. (2022). Emotion detection from text and sentiment analysis of Ukraine Russia war using machine learning technique. International Journal of Advanced Computer Science and Applications (IJACSA), 13(12).
4. Bandyopadhayay, D. K., Kadam, V., Aggarwal, K. K., Bhoite, U., Surendran, K., Ganesan, S., ... & Jain, M. V. (2012). Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM).
5. Guo, J. (2022). Deep learning approach to text analysis for human emotion detection from big data. Journal of Intelligent Systems, 31(1), 113-126.
6. Yan, Q. (2022). Real-Time Analysis of Youth Emotion Based on Python Language and Smart Sensor Network. Mobile Information Systems, 2022(1), 8635787.
7. Houssein, E. H., Hammad, A., & Ali, A. A. (2022). Human emotion recognition from EEG-based brain–computer interface using machine learning: a comprehensive review. Neural Computing and Applications, 34(15), 12527-12557.
8. Bota, P. J., Wang, C., Fred, A. L., & Da Silva, H. P. (2019). A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals. IEEE access, 7, 140990-141020.
9. Yang, L., Li, Y., Wang, J., & Sherratt, R. S. (2020). Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. IEEE access, 8, 23522-23530.
10. Alhalaseh, R., & Alasasfeh, S. (2020). Machine-learning-based emotion recognition system using EEG signals. Computers, 9(4), 95.
11. Ain, Q. T., Ali, M., Riaz, A., Noureen, A., Kamran, M., Hayat, B., & Rehman, A. (2017). Sentiment analysis using deep learning techniques: a review. International Journal of Advanced Computer Science and Applications, 8(6).
12. Casale, S., Russo, A., Scebba, G., & Serrano, S. (2008, August). Speech emotion classification using machine learning algorithms. In 2008 IEEE international conference on semantic computing (pp. 158-165). IEEE.
13. Doma, V., & Pirouz, M. (2020). A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals. Journal of Big Data, 7(1), 18.
14. Machová, K., Szabóova, M., Paralič, J., & Mičko, J. (2023). Detection of emotion by text analysis using machine learning. Frontiers in Psychology, 14, 1190326.