DAILY: AN AI POWERED DIARY APPLICATION WITH IMAGE GENERATION

Team Members:

Fatih Kapiz 202011064

Onurcan Erenel 202011071

Ahmet Buğra Yaka 201911068

Ümit Mete Şahin 202011021

Mehmet Mert Türkmen 201911064

Advisor:

Asst. Prof. Dr. Abdül Kadir Görür

Abstract

In the rapidly evolving area of Artificial Intelligence, its integration into our daily lives has become an important aspect. This paper focuses on enhancing our diary application by the usage of artificial intelligence, diving into topics such as emotion recognition. Through algorithm evaluation, we aim to optimize functionality. The focus lies on using Natural Language Processor (NLP) for emotion detection, to ensure accurate predictions. In this paper, we evaluate different approaches for this process. The objective is to select the most suitable AI model(s) for our application. This work aims to provide users with a deeper understanding of their emotions, revolutionizing the diary-writing experience.

Öz

Yapay zekânın hızla gelişen alanında, günlük hayatımıza entegrasyonu önemli bir husus haline gelmiştir. Bu makale, yapay zekânın kullanımıyla günlük uygulamamızı iyileştirmeye odaklanarak duygu tanıma gibi konuları derinlemesine incelemektedir. Algoritma değerlendirme yoluyla işlevselliği optimize etmeyi hedefliyoruz. Doğal Dil İşlemcisi'ni (NLP) duygu algılama için kullanarak doğru tahminler sağlamaya odaklanıyoruz. Bu makalede bu süreç için farklı yaklaşımları değerlendiriyoruz. Amacımız, uygulamamız için en uygun AI model(ler)ini seçmektir. Bu çalışma, kullanıcılara duygularını daha derinlemesine anlamalarını sağlamayı ve günlük tutma deneyimini devrimselleştirmeyi amaçlamaktadır.

Introduction

As the field of Artificial Intelligence continues to advance, its integration into our daily lives and digital interactions assumes an increasingly pivotal role. The primary objective of this paper is to enhance the interactivity of our diary application through the incorporation of Artificial Intelligence. In the course of this research, we explore prospective applications of Artificial Intelligence technologies, including text-based emotion recognition, text summarization, and generative image processing. Moreover, we delve into the evaluation of various algorithms, assessing their efficacy, and identify how we can employ these algorithms to serve our specific objectives.

One of the main features of our application is that the entries written by the user are going to be processed by our NLP model for emotion detection. To achieve correct and computationally efficient predictions we must analyze certain detection approaches, such as the rule-based approaches, machine learning approaches and the hybrid approach, which employs both of the machine learning and rule-based approaches. We also research generative image processing, and generate images based on the entries processed by our NLP engine. The application will exclusively accommodate the English language.

In this paper we survey through various AI models that are used for emotion detection in search of the most appropriate model or models for our application.

Natural Language Processor

Emotion Detection

Emotion detection starts with emotions. There are various forms of representing emotions [1]. Two main categories of emotion models are discrete and dimensional models. Some most widely used discrete emotional models are Ekman's six basic emotions [2] and Plutchik model which names eight fundamental emotions in opposite pairs [3]. Such as joy vs. sadness. Discrete models place emotions into distinct classes.

Dimensional emotion model is a theory that proposes that emotions can be represented by their position in a space defined by two or more dimensions. There are some models that represent emotions in 3 dimensions and 2 dimensions. Such as Russell and Mehrabian's 3 dimensional model [4] and Plutchik's 2 dimensional wheel of emotions [3]. While discrete models are more widely used, dimensional models better represent emotions that are similar to each other [1].

Choosing an emotional model also has to do with the detection algorithm that is going to be used and the dataset trained on. We have found several ready to use datasets, some of them represent emotions dimensionally while some represent them in a discrete way.

Detection Approaches

Engineering a natural language processor starts with selecting the tasks of the NLP model. We expect these tasks to change in time as our application grows in size, supports new features and pivots and evolves from the starting point, but we are going to start our research with an initial thought on these objectives and design choices we can take to overcome them.

We expect our NLP model to detect emotions, summarize content and rate the quality of posts, or filter out spam content. Detecting emotions requires rather less work compared to summarizing content.

Naive Bayes Algorithm

A supervised machine learning approach utilizing the Naive Bayes algorithm. Naive Bayes is a text classification algorithm that utilizes the Bayes' Theorem, which is a theorem regarding statistics and probability. The algorithm calculates the probability of an input belonging to a certain emotional category by calculating the probability based on a training dataset containing the probabilities[5].

Word Embeddings Method

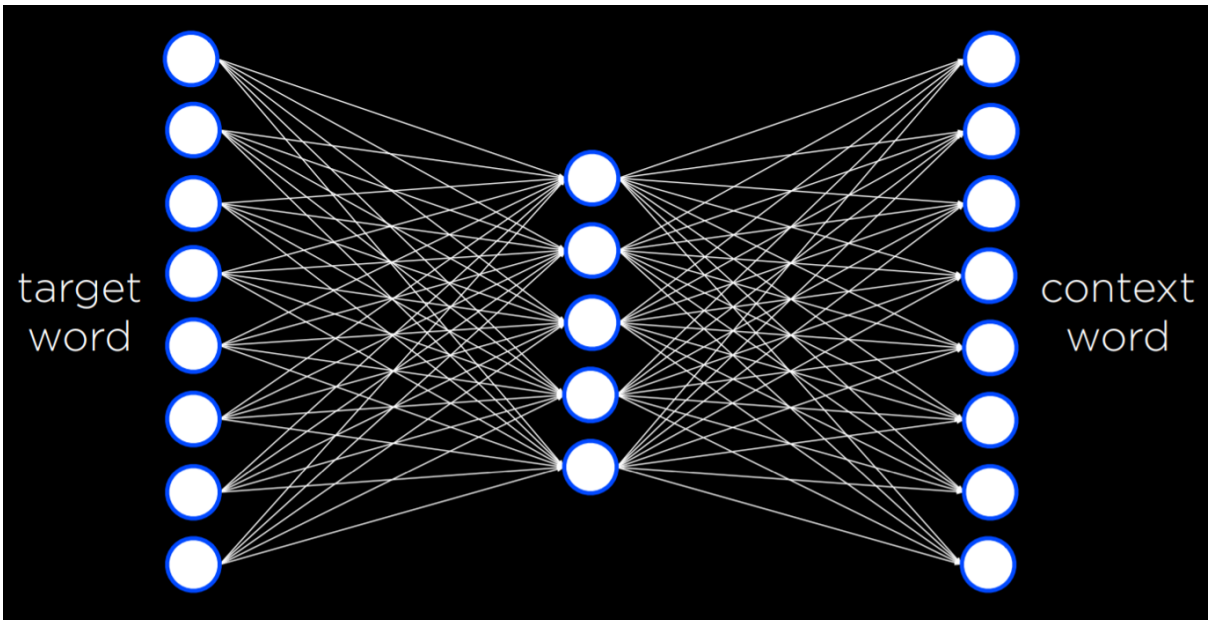

Figure 1 - Word2Vec Model [6]

Word2Vec is a neural network model that uses the Word Embeddings method to represent each word with a vector. One important aspect of this method is, now that we have vectors representing our words, we can train neural networks, which work well with vectors and numbers.

Word2Vec model learns by parameter optimization, using a backpropagation algorithm. The model adjusts the vector representations in each iteration of the training process to find a more accurate vector representation of words. There is no need to code a Word2Vec from the start-up in our case. The Word2Vec model exists within the **Gensim** library in Python and can be given tokenized texts to train and can be used together with the **Scikit-learn** library with ease.

Some use cases of this method could be to generate words that associate most with a certain word. We can check this association just by a distance function. Running this distance function for each word in pairs with our certain word then selecting the closest distance pairs we can arrive at the associations and extract semantic meanings.

Lilleberg, Zhu and Zhang achieved a successful prediction rate of +70% in their work utilizing a Word2Vec model [7].

Transformer Based Models for Emotion Detection

For us to understand the Transformer models, we need to understand what is an **attention mechanism** and what is a **self-attention mechanism**. An **attention mechanism** is a mechanism that enables neural networks to focus on the features of the input data dynamically [8]. A **self-attention mechanism** involves an attention mechanism that considers various positions within a single sequence in order to generate a representation for that sequence [9]. Transformer models are using self-attention mechanisms to compute representations of its input [9]. This enables them to effectively process and understand context-rich information.

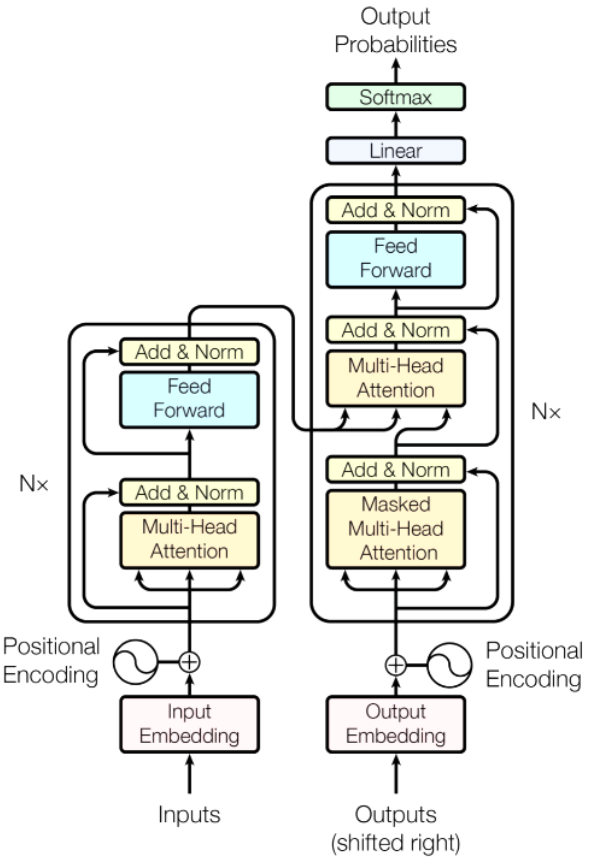

Figure 2 - Transformer Model Architecture [9]

In the context of emotion detection, transformer-based models such as **BERT**, **GPT**, and **RoBERTa** have demonstrated remarkable performance [10]–[12]. These models can be fine-tuned specifically for emotion detection tasks, leveraging their pre-trained knowledge to effectively identify and classify emotions in text [13].

Pros and Cons

1. Naive Bayes Algorithm:
    - Pros:
        - Simple and easy to implement.
        - Works well in small training datasets.
        - Efficient for text classification cases.
    - Cons:
        - Assumes that features are independent.
        - Cannot capture the complex relationships between the words.
2. Word Embeddings Method (Word2Vec):
    - Pros:
        - Represents words in a vector, encapsulating semantic relationships.
        - Can be pre-trained on large datasets.
    - Cons:
        - For best results, a great amount of data is needed.
        - Cannot capture context-specific details of words.
3. Transformer Based Models:
    - Pros:
        - Can capture context-specific details.
        - Cutting edge results.
    - Cons:
        - Requires powerful hardware for training.
        - Large memory requirements.

Conclusion

This report highlights the significance of AI, particularly in emotion detection to enhance user experience in our diary application. We've examined the standard approaches, from Naive Bayes to more advanced models like Transformer models. Each offering some pros and some cons. As the field of AI continues to evolve, maintaining a keen awareness of growing technologies and upholding the ethical considerations is vital for our applications success and impact.

References

[1] F. A. Acheampong, C. Wenyu, and H. Nunoo-Mensah, “Text-based emotion detection: Advances, challenges, and opportunities,” *Eng. Rep.*, vol. 2, no. 7, p. e12189, 2020, doi: 10.1002/eng2.12189.

[2] A. Seyeditabari, N. Tabari, and W. Zadrozny, “Emotion Detection in Text: a Review.” arXiv, Jun. 02, 2018. doi: 10.48550/arXiv.1806.00674.

[3] R. Plutchik, “Chapter 1 - A GENERAL PSYCHOEVOLUTIONARY THEORY OF EMOTION,” in *Theories of Emotion*, R. Plutchik and H. Kellerman, Eds., Academic Press, 1980, pp. 3–33. doi: 10.1016/B978-0-12-558701-3.50007-7.

[4] J. A. Russell and A. Mehrabian, “Evidence for a three-factor theory of emotions,” *J. Res. Personal.*, vol. 11, no. 3, pp. 273–294, Sep. 1977, doi: 10.1016/0092-6566(77)90037-X.

[5] “What are Naive Bayes classifiers? | IBM.” Accessed: Nov. 06, 2023. [Online]. Available: https://www.ibm.com/topics/naive-bayes

[6] “Lecture 6 - CS50’s Introduction to Artificial Intelligence with Python.” Accessed: Nov. 07, 2023. [Online]. Available: https://cs50.harvard.edu/ai/2023/notes/6/

[7] J. Lilleberg, Y. Zhu, and Y. Zhang, “Support vector machines and Word2vec for text classification with semantic features,” in *2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)*, Jul. 2015, pp. 136–140. doi: 10.1109/ICCI-CC.2015.7259377.

[8] A. Galassi, M. Lippi, and P. Torroni, “Attention in Natural Language Processing,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 10, pp. 4291–4308, Oct. 2021, doi: 10.1109/TNNLS.2020.3019893.

[9] A. Vaswani *et al.*, “Attention Is All You Need.” arXiv, Aug. 01, 2023. doi: 10.48550/arXiv.1706.03762.

[10] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” arXiv, May 24, 2019. doi: 10.48550/arXiv.1810.04805.

[11] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, “Improving Language Understanding by Generative Pre-Training”.

[12] “RoBERTa: An optimized method for pretraining self-supervised NLP systems.” Accessed: Nov. 07, 2023. [Online]. Available: https://ai.meta.com/blog/roberta-an-optimized-method-for-pretraining-self-supervised-nlp-systems/

[13] F. Acheampong, H. Nunoo-Mensah, and W. Chen, *Transformer Models for Text-based Emotion Detection: A Review of BERT-based Approaches*. 2021.