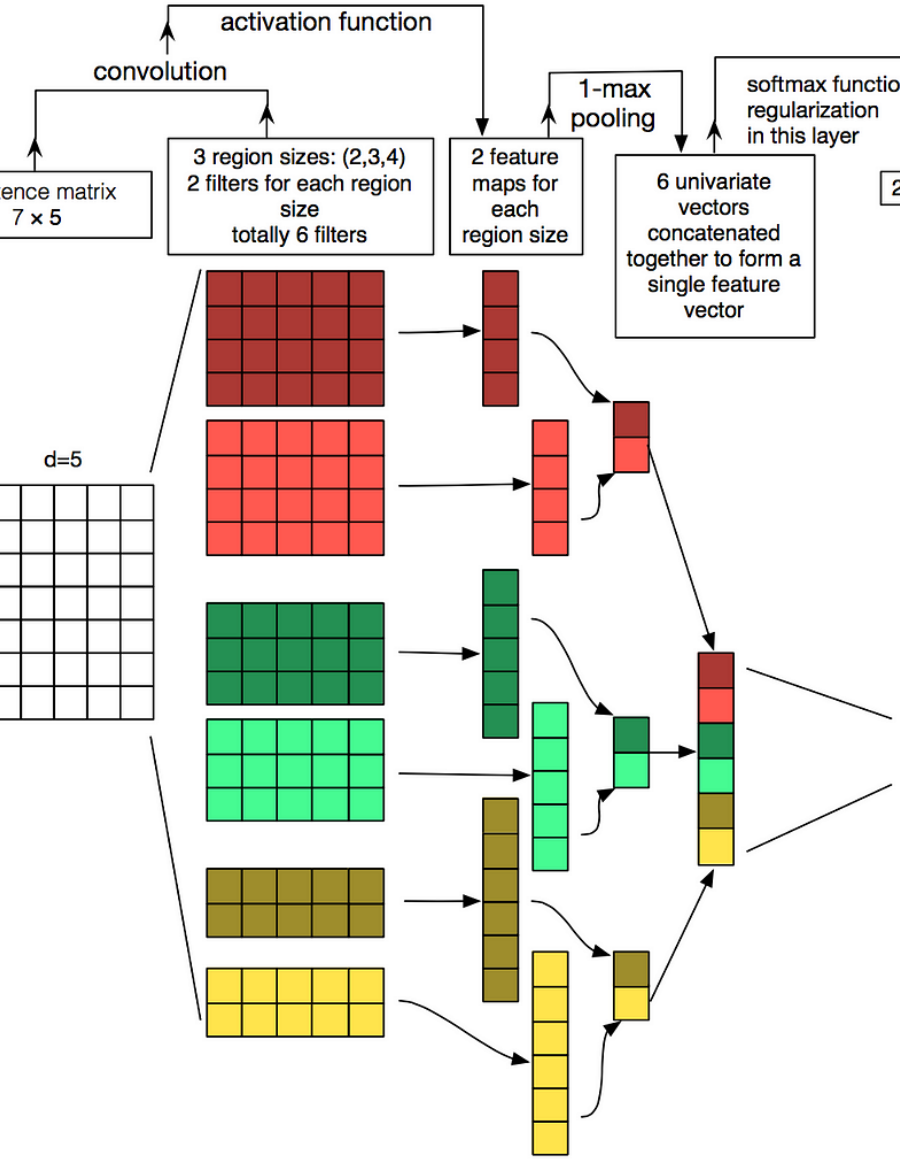


Text Classification with CNN and Entity Extraction with NER

In this presentation, we will explore the power of text classification using Convolutional Neural Networks (CNN) and entity extraction using Named Entity Recognition (NER). These techniques have revolutionized the field of natural language processing and have numerous real-world applications.

by Mitul Varshney

Introduction to Text Classification

What is Text Classification?

Text classification is the process of assigning predefined categories to textual data. It can be used for sentiment analysis, spam filtering, and more.

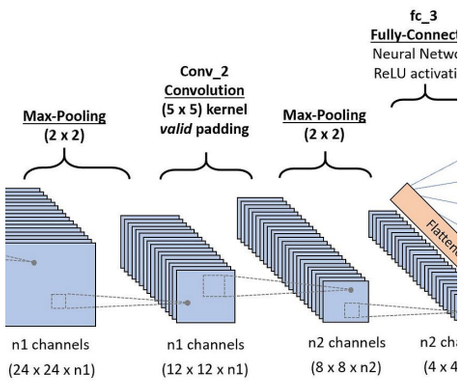
Why is it Important?

Text classification allows us to process and analyze large amounts of unstructured data. It helps us to understand consumer sentiment, predict trends, and make more informed decisions.

How Does it Work?

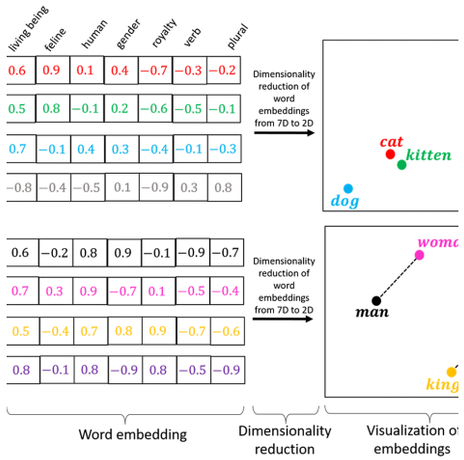
Text classification algorithms use machine learning models to analyze textual data. These models are trained on labeled datasets to recognize patterns and predict categories.

Convolutional Neural Networks (CNN)



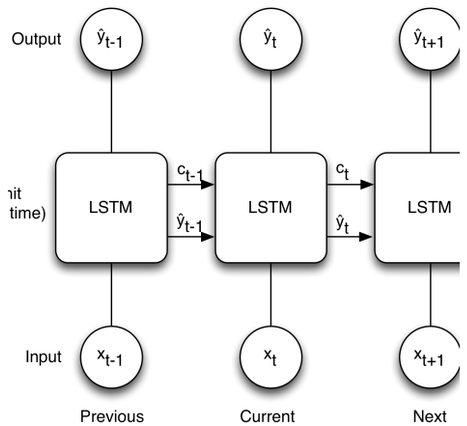
How CNN works in Text Classification

CNNs use a layered approach to analyze text. Each layer processes the text and extracts increasingly complex features. This allows the network to recognize relevant patterns and make accurate predictions.



How text is converted to numeric form

Before being processed by a CNN, text is converted to a numeric form using techniques like word embedding. This allows the text to be fed into the network as input.



Advantages and Limitations of CNN

CNNs are extremely effective at processing large amounts of text data and achieving high accuracy. However, they can be resource-intensive and may suffer from overfitting if not trained properly.

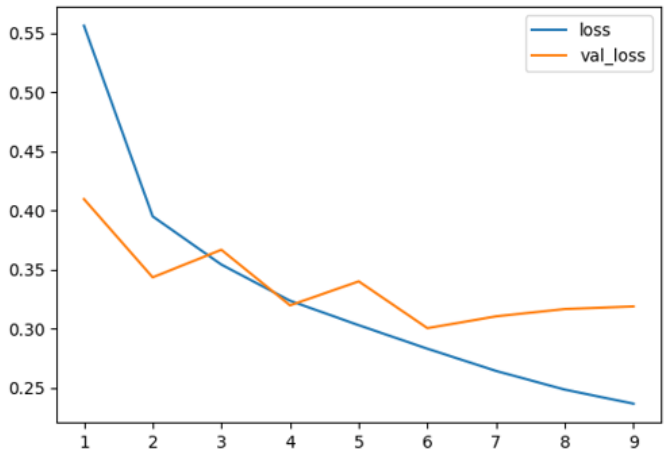
Results of training our CNN model

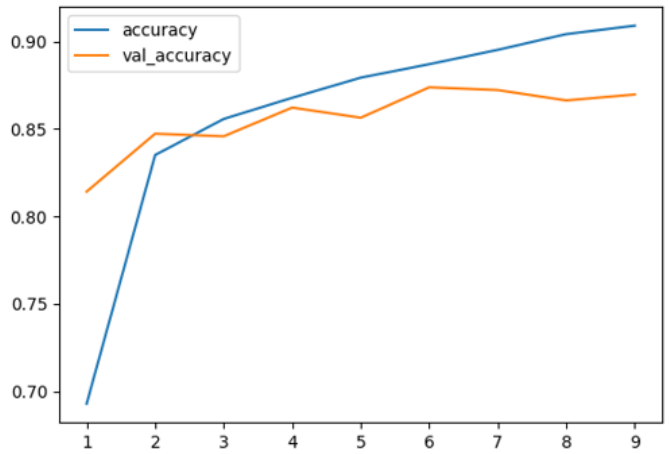
Our model was trained on the Large Movie Dataset from IMDb to classify human text into two classes:

- positive - It was a positive text, good for morning news. e.g. Prime Minister visited kids
- negative - It wasn't a particularly good news, keep it for later. e.g. Terrorist attack

Graphs for loss and accuracy

The model has reached an validation accuracy of **85%** within **9** epochs using an early stopping callback.





Entity Extraction using Named Entity Recognition (NER)

In the 19th century, there was something called the “cult of domesticity” for many American women. This meant that most married women were expected to stay in the home and raise children. As in other countries, American wives were very much under the control of their husband, and had almost no rights. Women who were not married had only a few jobs open to them, such as working in clothing factories and serving as maids. By the 19th century, women such as Lucretia Mott and Elizabeth Cady Stanton thought that women should have more rights. In 1848, many of these women met and agreed to fight for more rights for women, including voting. Many of the women involved in the movement for women’s rights were also involved in the movement to end slavery.

WIKIPEDIA

Tag colors:

LOCATION PERSON TERM DATE CONDITION PROCESS PEOPLE

What is NER and How it Works

NER is the process of identifying and extracting entities, such as names, locations, and dates, from unstructured text. It is achieved using machine learning algorithms that are trained on labeled datasets.

Advantages and Limitations of NER

NER can help to extract useful information from large amounts of text data. It is especially useful for tasks like information retrieval and data mining. However, NER algorithms can struggle with identifying ambiguous or uncommon entities.

Combined Approach of CNN and NER

Benefits of Using Both Methods

Using both CNN and NER allows us to achieve greater accuracy and more nuanced analysis. This can lead to more accurate sentiment analysis, more precise topic modeling, and more.

1

How to use both methods together

By combining CNN and NER, we can achieve even better results at text classification. NER can help to identify relevant entities, while CNN can analyze the text around those entities to extract deeper meaning.

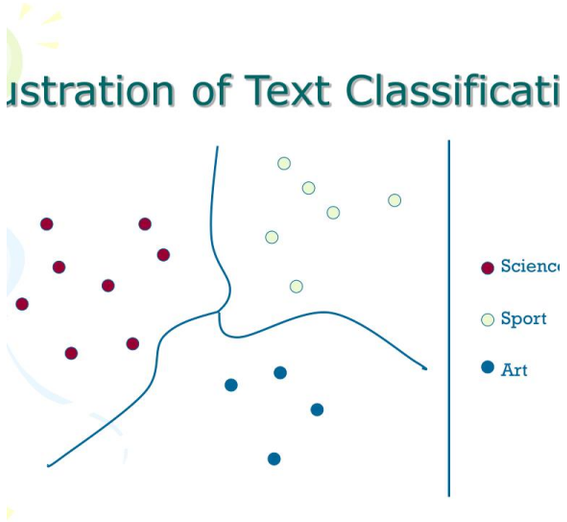
2

3

Examples of combined approach

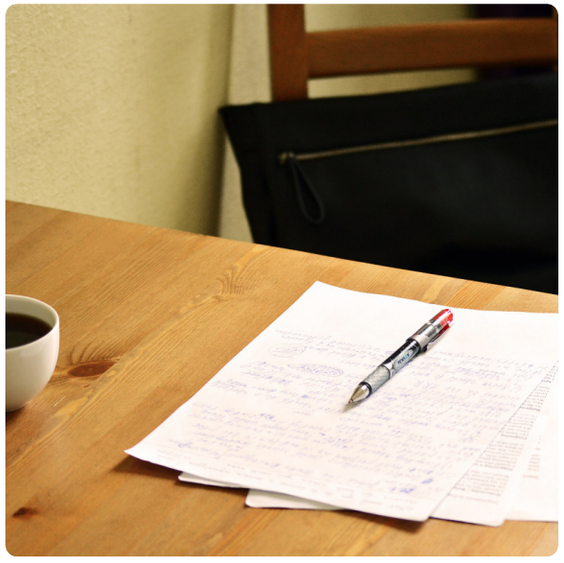
The combined approach has been used in a range of applications, from analyzing social media feeds to identifying risk factors in medical records.

Conclusion and Real-world Applications of Text Classification with CNN and NER



Real-World Applications

The combined approach of CNN and NER has numerous applications, from fraud detection to recommendation systems. It can help organizations to better understand their customers, identify trends, and make more informed decisions.



Conclusion

In conclusion, text classification with CNN and entity extraction with NER are powerful tools for analyzing unstructured text data. By combining these techniques, we can achieve even better results and gain deeper insights into the meaning behind the text.