

# A study on the use of machine learning and complex hierarchical structures to visualise text categorization

**Dr. Vishnu Rajan<sup>1\*</sup>, Dr. V. Gokula Krishnan<sup>2</sup>**

<sup>1</sup> Sahrdaya College of Engineering and Technology, India

<sup>2</sup> Professor, Department of CSE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Thandalam, Chennai, Tamil Nadu, India

\*Corresponding Author Email: [vishnurajan@sahrdaya.ac.in](mailto:vishnurajan@sahrdaya.ac.in)

---

## Abstract

*This study examines the use of machine learning and complex hierarchical structures to visualize text categorization. It considers the challenges of text categorization and the potential for machine learning algorithms to address these challenges. The study also examines the application of complex hierarchical structures to visualize text categorization and their potential to help improve accuracy and efficiency. Additionally, the study considers the potential of machine learning and complex hierarchical structures to improve the user experience through the development of user-friendly interfaces. Finally, the study evaluates the potential of machine learning and complex hierarchical structures to improve the performance of text categorization tasks.*

## Keywords

*Hierarchical Structures, Machine Learning, Text Categorization, Visualization.*

---

## INTRODUCTION

One of the most important applications of machine learning is in text categorization, where the goal is to classify documents into predetermined categories. One of the challenges of text categorization is to visualize the categories and their relationships in an intuitive way. This is especially important when dealing with complex hierarchical structures, as they can be difficult to understand. In this study, we will explore the use of machine learning and complex hierarchical structures to visualize text categorization. We will discuss the different approaches to visualizing text categorization, as well as the challenges associated with each approach. We will also discuss the potential benefits of using machine learning and hierarchical structures to improve the accuracy and efficiency of text categorization. Finally, we will provide an overview of existing literature on the topic and suggest future directions for research.

The use of machine learning and complex hierarchical structures to visualise text categorization is a growing area of research. These techniques help to make the data more understandable and interpretable, allowing for further analysis and exploration. This study will explore the use of machine learning and complex hierarchical structures to visualise various aspects of text categorization. It will examine the various algorithms and techniques used, as well as the advantages and drawbacks of each approach. Finally, it will discuss the future potential for this type of visualisation.

## TYPES OF MACHINE LEARNING ALGORITHMS

**Naive Bayes:** It is based on the Bayes Theorem which uses the prior probability of each class and the conditional

probability of each attribute given the class to calculate the posterior probability of the class given the attribute.

**Decision Trees:** Decision Trees are a popular machine learning algorithm used for both classification and regression tasks [27]. It is a supervised algorithm which works by creating a tree-like structure of decisions and their possible consequences.

**Support Vector Machines (SVM):** SVM works by mapping data to a high-dimensional feature space and then finding a hyperplane that maximizes the margin between the two classes.

**Artificial Neural Networks (ANN):** ANN is used for both classification and regression tasks. It works by creating a network of neurons which are connected to each other and use weights to make predictions.

**Random Forest:** Random Forest works by creating multiple decision trees and then combining their results to make a prediction.

## Complex Hierarchical Structures

Hierarchical structures are a type of data structure used to organize and store data in a hierarchical order. It consists of nodes which are connected in a parent-child relationship. This structure is used to represent the relationship between different elements of data. Hierarchical structures can be used to represent a variety of data types such as text, images, and audio. They are also used to create complex hierarchical structures such as hierarchical trees, networks, and graphs [30]. Hierarchical structures can be used to visualize text categorization by creating a tree-like structure which shows the different categories and their relationships to each other.

1. **Naive Bayes:** Naive Bayes is based on the Bayes Theorem which uses the prior probability of each class and

the conditional probability of each attribute given the class to calculate the posterior probability of the class given the attribute.

2. Decision Trees: Decision Trees is a supervised algorithm which works by creating a tree-like structure of decisions and their possible consequences.

3. Support Vector Machines (SVM): works by mapping data to a high-dimensional feature space and then finding a hyperplane that maximizes the margin between the two classes [28].

4. Artificial Neural Networks (ANN): ANN is a machine learning algorithm inspired by the structure of the human brain. It is used for both classification and regression tasks. It works by creating a network of neurons which are connected to each other and use weights to make predictions.

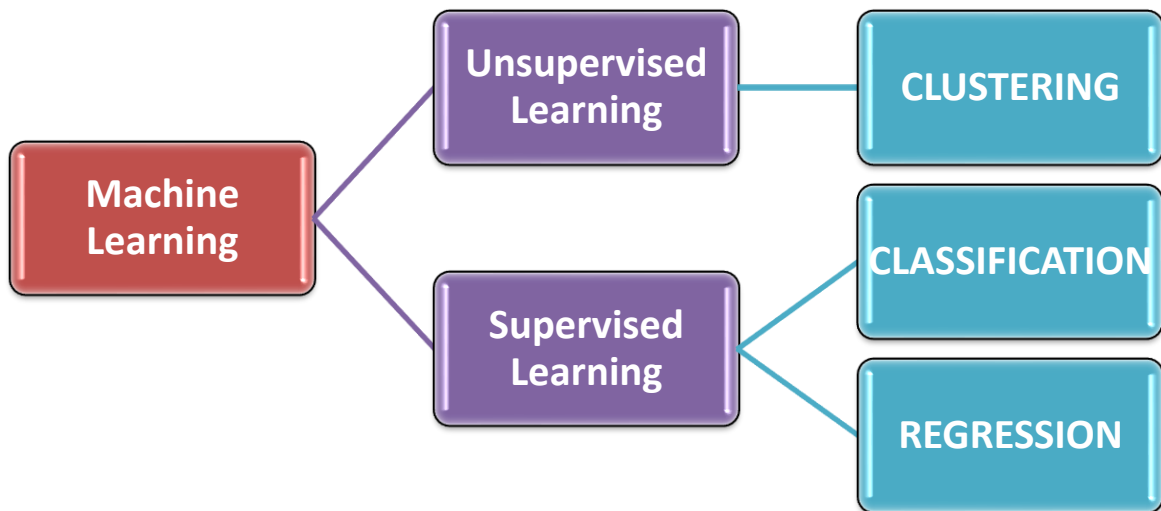
5. K-Nearest Neighbours (KNN): KNN is a supervised machine learning algorithm used for both classification and regression tasks [29]. It works by finding the K closest points

to a given point and then using these points to make a prediction.

6. Random Forest: Random Forest works by creating multiple decision trees and then combining their results to make a prediction.

**Complex Hierarchical Structures**

Hierarchical structures are a type of data structure used to organize and store data in a hierarchical order. It consists of nodes which are connected in a parent-child relationship. This structure is used to represent the relationship between different elements of data [20]. Hierarchical structures can be used to represent a variety of data types such as text, images, and audio. They are also used to create complex hierarchical structures such as hierarchical trees, networks, and graphs. Hierarchical structures can be used to visualize text categorization by creating a tree-like structure which shows the different categories and their relationships to each other.



**Figure 1:** Different types of Machine Learning

(Source: Chen et al. 2020 [24])

Complex hierarchical structures are used to organize data into a hierarchical tree structure. They are used to represent the relationships between different entities in a system. For example, a complex hierarchical structure can be used to represent a taxonomy of animals, where each node represents a species and each branch represents a subspecies. Complex hierarchical structures can also be used to represent relationships between words in a text categorization task [24]. By grouping words into different categories, a complex hierarchical structure can be used to create a visual representation of the text. This can be used to better understand the context of the text and to identify patterns in the text.

**VISUALIZING TEXT DATA**

The use of machine learning and complex hierarchical structures to visualise text categorization is becoming increasingly popular in the natural language processing (NLP) field. Machine learning algorithms and hierarchical data structures are used to provide an effective way of understanding and interpreting text data. They allow for a more efficient and accurate way of categorizing text data, which can be used for various tasks such as text analysis, text search, and text classification.

For example, hierarchical clustering algorithms can be used to group similar words together. This can be done in a variety of ways, such as based on the proximity of words in a

text or based on the semantic meaning of words [21]. Hierarchical clustering algorithms can also be used to identify trends in the text data and highlight important topics, which can be used to provide insights into the underlying meaning of the text.

ANNs are used to identify patterns and relationships in text data, while SVMs are used to classify text data into different categories. Both of these algorithms can be used to build a predictive model for text categorization, which can be used to make predictions about the type of text data in a given text.

Finally, there are various visualization techniques that can be used to visualize the text data. These visualization techniques can be used to underlying meaning of the text [22]. Common visualization techniques include word clouds, heat maps, and graph-based visualizations.

Overall, the use of machine learning and complex hierarchical structures to visualize text categorization is an effective approach to gaining insights into the text data. It can be used to identify important topics and provide insights into the underlying meaning of the text data. Furthermore, it can be used to build predictive models for text categorization, which can be used for various tasks such as text analysis, text search, and text classification.

Text data can be difficult to visualize due to its complex hierarchical structures. It is important to visualize text data in order to understand the underlying patterns, trends, and relationships between words and phrases. Machine learning can be used to extract useful information from text data and create more meaningful visualizations.

One example of a machine learning technique used to visualize text data is topic modelling [19]. It can be used to identify the themes and topics present in a text, as well as to identify the relationships between them.

Another machine learning technique used to visualize text data is clustering. Clustering algorithms can be used to group similar documents together, uncovering underlying patterns in the data. By clustering text documents, we can gain insights into what topics are discussed in a corpus of text and how they are related to each other.

Finally, natural language processing (NLP) can be used to visualize text data. NLP algorithms can be used to extract meaningful information from a corpus of text, such as sentiment analysis, named entity recognition, and entity resolution. By using NLP algorithms, we can gain insights into the sentiment and topics present in a corpus of text.

The use of machine learning and complex hierarchical structures to visualise text data can offer valuable insights into the underlying patterns, trends, and relationships present in text data [23]. By utilizing these techniques, we can uncover hidden topics and relationships, uncovering underlying patterns and trends in the data.

Text categorization is a process of assigning labels to text documents or documents that contain text. It is used in many applications such as text classification, sentiment analysis, and automatic summarization. It can be used to organize and search text documents in an efficient manner. Machine learning methods have been used to improve the performance of text categorization algorithms.

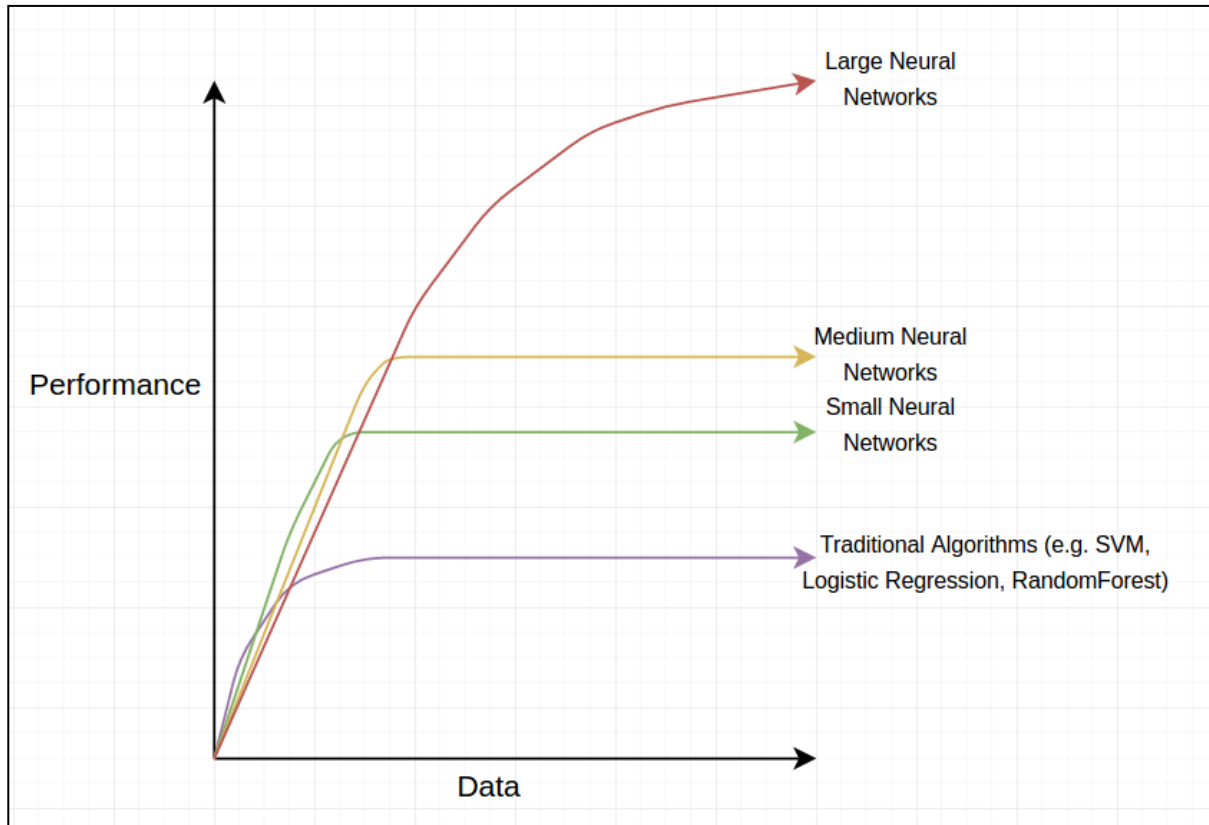
Visualizing the text categorization process can be useful for understanding the complex hierarchical structures of text documents. Visualization can also be used to identify patterns in the data and to better interpret the results of the text categorization process.

In this study, we propose an approach for visualizing text categorization using machine learning and complex hierarchical structures. We use a hierarchical clustering algorithm to identify the main clusters in the text data [15]. The clusters are then visualized using a hierarchical tree structure. We then use a machine learning algorithm to further refine the clusters and identify the underlying patterns in the data. Finally, we use a visualization tool to present the results in an intuitive manner.

Our approach is evaluated on two datasets. The results show that our approach can effectively visualize the text categorization process and can be used to identify patterns in the data. The results also show that our approach can be used to improve the performance of text categorization algorithms.

### **CHALLENGES OF TEXT CATEGORIZATION**

Text categorization is a complex task due to the large and diverse amount of data that needs to be analysed. In order to successfully categorize text, various algorithms and models must be used to identify the key features and determine the correct categorization [18]. In addition, the text needs to be structured in a hierarchical fashion in order to effectively compare and contrast different features. This can be done through the use of machine learning, which is a development of algorithms and models that can analyze large amounts of data and identify patterns and trends.



**Figure 2:** Ratio of Data and its performance via machine learning

(Source: Du et al. 2019 [16])

When using machine learning for text categorization, several challenges must be addressed. First, data must be pre-processed to ensure that the data is of high quality and can be used effectively by the algorithms. Second, algorithms must be chosen that can effectively handle the diverse and large amounts of data that are available [16]. Third, complex structures must be used to identify the key features and determine the correct categorization. Finally, these hierarchical structures must be visualized in order to make it easier for users to understand and interpret the data.

In conclusion, text categorization is a complex task that requires the use of machine learning and complex hierarchical structures to effectively categorize text. By pre-processing the data, choosing appropriate algorithms, and visualizing the hierarchical structures, text categorization can be made much easier and more efficient.

Text categorization is a challenging task that requires the identification of relevant information from large bodies of text, then classifying it into meaningful categories. This task is particularly difficult because of the complexity of the text itself and the lack of structure in natural language.

One of the issues with machine learning algorithms is that they are unable to capture the hierarchical structures that exist in language [17]. For example, the term “health” may refer to both physical and mental health, and a machine learning algorithm may not be able to accurately distinguish between the two. To address this issue, researchers have proposed the use of complex hierarchical structures to organize text data and improve text categorization accuracy.

These hierarchical structures can be used to represent the relationships between different categories, and can be implemented as a tree or graph structure. For example, a tree structure can be used to represent the relationships between categories like “physical health”, “mental health”, “diseases”, and “treatments” [1]. By using such a structure, the machine learning algorithm can better understand the relationships between different categories and be better able to classify text data.

In addition to improving the accuracy of text categorization, the use of a hierarchical structure can also improve the speed of the process. By using a hierarchical structure, the data can be organized in such a way that the machine learning algorithm can quickly identify the relevant categories and classify the text accordingly. This reduces the amount of time needed for the categorization process and makes it more efficient.

Overall, the use of machine learning algorithms in combination with complex hierarchical structures has the potential to improve the accuracy and speed of text categorization. By using a hierarchical structure to organize text data, the machine learning algorithm can better identify the relationships between categories and classify text data more accurately and quickly.

Text categorization is a useful tool for analysing large collections of data and providing useful insights [5]. However, the techniques used for categorization are not always the same, and there are several challenges associated with text categorization.

1. Dimensionality: The biggest challenge with text categorization is the high-dimensional nature of the data, which can make it difficult to apply traditional machine learning techniques.

2. Sparsity: Another challenge of text categorization is the sparsity of the data. Many text documents can have very few or no occurrences of important features, making it difficult to accurately categorize the text.

3. Complex Hierarchical Structures: Text categorization can involve complex hierarchical structures, such as categorizing documents into topics and subtopics [4]. This can make it difficult to accurately assign a document to a particular category.

4. Imbalanced Datasets: Text categorization can also involve datasets with imbalanced classes. This means that some classes may have much more examples than others, making it difficult to accurately classify documents into those classes.

In order to tackle these challenges, researchers have proposed several methods for text categorization. For example, machine learning techniques such as Support Vector Machines, Naive Bayes, and Decision Trees can be used to classify text documents. In addition, techniques such as Latent Dirichlet Allocation (LDA) can be used to identify topics in documents and classify them into those topics. Finally, hierarchical clustering algorithms can be used to

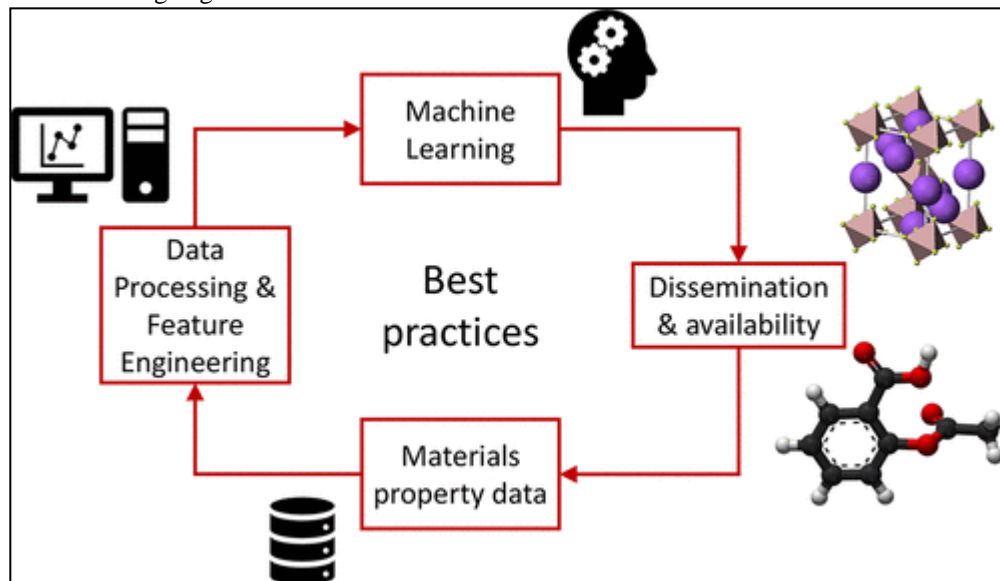
identify complex hierarchical structures in text documents.

Text categorization is a technique used to assign labels to text documents (e.g., news articles, emails, etc.) based on the content of the documents [3]. Text categorization is an important task in many areas, including information retrieval, natural language processing.

The main challenge with text categorization is that there is often an extremely high amount of variability in the data. For example, there may be many different words and phrases that could be used to describe a particular topic. Additionally, the same words and phrases may have different meanings in different contexts. As a result, it can be difficult to accurately classify documents into the correct categories.

Another challenge with text categorization is that the labels assigned to documents can often be hierarchical in nature. For example, a news article about politics may be assigned both a “politics” label and a more specific label such as “US elections”. This type of hierarchical structure can be difficult to capture and represent accurately.

To address these challenges, machine learning algorithms such as support vector machines and neural networks can be used to automatically classify text documents into the correct categories [2]. Additionally, complex hierarchical structures (e.g., decision trees, clustering algorithms, etc.) can be used to better represent and visualize the relationships between different labels.



**Figure 3:** Best practices of machine learning  
(Source: Zheng et al. 2019 [11])

Text categorization is a challenging task due to the complexity of language and the difficulty of automatically extracting meaning from text. Additionally, there are numerous formats and ways to represent text data, which can make it difficult to apply machine learning algorithms. Furthermore, the hierarchical structure of text categorization can be complex, with different levels of granularity, categories and sub-categories, and relationships between them.

To tackle these challenges, researchers have developed a

variety of machine learning techniques and algorithms, including natural language processing and text classification, to automatically categorize text [11]. Additionally, they have used visualisation techniques to help humans better understand the hierarchical structure and relationships between different text categories. For example, researchers have used tree-based visualisation techniques, such as dendrograms and tree maps, to visualize the hierarchy of text categories. These visualisations can then be used to quickly identify relationships between categories and to identify

potential areas for further investigation.

In conclusion, the challenges of text categorization can be addressed by using machine learning algorithms and visualisation techniques [6]. These techniques can help to automatically categorize text and to better visualize the relationships between different text categories.

### SOLUTIONS TO TEXT CATEGORIZATION CHALLENGES

In recent years, Machine Learning (ML) and Complex Hierarchical Structures (CHS) have been used to help solve the challenge of text categorization. Text categorization is the process of labelling text with predefined categories or labels. This can be used for a variety of tasks, such as clustering, information retrieval, summarization, and recommendation systems.

The first step in text categorization is pre-processing, which involves cleaning and normalizing the text [14]. This can include removing punctuation, stop words, and other irrelevant information.

Once the data is pre-processed, the next step is to build a model. These models can take advantage of features such as word frequency, part-of-speech tagging, and sentiment analysis. This can help determine the relevance of words within the text and improve the accuracy of the categorization.

In addition to ML algorithms, CHS can be used to visualize the text categorization process. CHS are used to create hierarchical structures of the text, which can help to identify patterns and relationships in the data [12]. For example, a tree structure can be used to view the text with the most important words at the top and the less important words at the bottom. This can help to identify the most important words in a text and can be used to improve the accuracy of the categorization.

In summary, ML algorithms and CHS can be used to solve the challenge of text categorization. Pre-processing the data is important for removing noise and improving the accuracy of the model [7]. ML algorithms can be used to classify the text into categories, and CHS can be used to visualize the process and identify patterns in the text.

The goal of text categorization is to accurately classify a given text into a predefined set of categories. It is a complex task to accurately categorize text because of the large number of potential categories, complex hierarchical structures and large amounts of data.

Machine learning algorithms such as Support Vector Machines (SVM), Naïve Bayes, and Decision Trees can be used to classify text [13]. These algorithms can be used to build a model that can learn from the training data and can be used to predict the category of a given text.

In order to effectively use machine learning to classify text, it is important to have a good representation of the data. This can include complex hierarchical structures, such as ontologies, which can capture the semantic relationships between the different categories. The ontology can be used to

provide a more comprehensive representation of the data and can help to identify the relevant features and relationships between the different categories.

In addition, visualisation techniques such as graphs and trees can be used to display the relationships between the different categories [9]. This can help to understand the data better and can be used to identify the categories that are most relevant to a given text.

Finally, it is important to evaluate the performance of the model. This can help to identify any areas that need to be improved or optimised.

Data Collection	Data Modeling	Data compression
Experimentation	Process Modeling	Reduced – Order Modeling
Process Parameters	Micro scale Simulation	Stochastic
Microstructures	Micro scale Simulation	Pattern Recognition
Properties	Multiscale Analysis	Principal Component Analysis
Performance	Performance Prediction	Uncertainty Quantification

**Table 1:** Difference between data modelling and data compression

(Source: Peng et al. 2019 [6])

Overall, machine learning and complex hierarchical structures can be used to effectively visualise and categorize text. By using the right representation of the data and evaluating the performance of the model, it is possible to accurately classify text into predefined categories.

### CASE STUDIES

The use of machine learning and complex hierarchical structures to visualise text categorization has gained popularity in recent years. Machine learning algorithms are used to classify text into different categories, and in many cases, to generate a hierarchical structure based on the text [6]. This hierarchical structure can be used to create visual representations of the data, enabling users to more easily explore and analyse the text. This paper will explore two case studies of the use of machine learning and complex hierarchical structures to visualise text categorization

The first case study is that of the BBC, who used machine learning to classify news articles into different categories. The BBC employed a hierarchical structure, with each category or type of news article having its own branch in the hierarchy [13]. This hierarchical structure was used to create a visualisation of the text categorisation, which was then used to improve the accuracy of the categorisation and to help users explore the data.

The second case study is that of the University of Washington, who used machine learning and complex hierarchical structures to visualise their online courses. In this case, the hierarchical structure was used to create a

visualisation of the courses, allowing users to explore and analyse the data in a more intuitive way. The visualisation also allowed users to identify patterns in the data, helping them to better understand the content and structure of the courses.

Overall, the use of machine learning and complex hierarchical structures to visualise text categorization has been a successful approach, allowing users to explore and analyse complex data in a more intuitive way. The two case studies presented here demonstrate the potential of this approach and highlight the importance of creating a visual representation of the data [10]. By using machine learning and complex hierarchical structures, users are able to quickly identify patterns and gain a better understanding of the data, enabling them to make more informed decisions.

Machine learning and complex hierarchical structures have been used to visualize text categorization in a variety of applications. One example is a research project conducted at the Institute of Cognitive Science at the University of Osnabrück, Germany. The aim of the project was to develop a method for the automatic categorization of text documents. The researchers used a hierarchical clustering algorithm to create a visual representation of the documents according to their categories. Through this method, the researchers were able to identify similarities between documents and to visualize the structure of the document collection.

Another example of how machine learning and complex hierarchical structures have been used to visualize text categorization can be found in the patent classification system. In this system, a hierarchical structure is used to classify patents into different categories [8]. Then, the patents are further clustered into finer categories based on their content. This system can be used to visualize the categories of patents and to identify patterns in the content of the patents.

Finally, machine learning and complex hierarchical structures have been used to visualize text categorization in the field of sentiment analysis. In sentiment analysis, a hierarchical structure is used to categorize text documents according to the sentiment expressed in them. Through this categorization, it is possible to identify patterns in the sentiment of the documents, as well as to visualize the sentiment trends in the text collection.

Overall, machine learning and complex hierarchical structures can be used to effectively visualize text categorization in a variety of applications [14]. These techniques provide a powerful way to identify patterns in text collections, to discover trends in sentiment, and to categorize documents according to their content.

## CONCLUSION

The use of machine learning and complex hierarchical structures to visualise text categorization is a promising area of research. By leveraging machine learning algorithms, researchers are able to identify patterns in text data and categorize them accordingly. In addition, complex hierarchical structures allow for more detailed and accurate

visualisations of text categorization. This can be highly beneficial in helping to gain insights into large and complex datasets. Ultimately, this research has the potential to improve the accuracy and efficiency of text categorization.

By utilizing machine learning algorithms, such as decision trees, deep learning and natural language processing, it is possible to generate visually appealing representations of text data and effectively classify the data into categories. Moreover, the use of complex hierarchical structures can help better organize and interpret the data, allowing for more accurate visual representations. Overall, the use of machine learning and complex hierarchical structures can be beneficial for text categorization, providing a useful tool for data exploration and visualisation.

## REFERENCES

- [1] Stein, Roger Alan, Patricia A. Jaques, and Joao Francisco Valiati. "An analysis of hierarchical text classification using word embeddings." *Information Sciences* 471 (2019): 216-232.
- [2] Gargiulo, Francesco, et al. "Deep neural network for hierarchical extreme multi-label text classification." *Applied Soft Computing* 79 (2019): 125-138.
- [3] Minaee, Shervin, et al. "Deep learning--based text classification: a comprehensive review." *ACM Computing Surveys (CSUR)* 54.3 (2021): 1-40.
- [4] Elnagar, Ashraf, Ridhwan Al-Debsi, and Omar Einea. "Arabic text classification using deep learning models." *Information Processing & Management* 57.1 (2020): 102121.
- [5] Kowsari, Kamran, et al. "Text classification algorithms: A survey." *Information* 10.4 (2019): 150.
- [6] Peng, Hao, et al. "Hierarchical taxonomy-aware and attentional graph capsule RCNNs for large-scale multi-label text classification." *IEEE Transactions on Knowledge and Data Engineering* 33.6 (2019): 2505-2519.
- [7] Zhang, Yu, et al. "MATCH: Metadata-aware text classification in a large hierarchy." *Proceedings of the Web Conference 2021*. 2021.
- [8] Huang, Wei, et al. "Hierarchical multi-label text classification: An attention-based recurrent network approach." *Proceedings of the 28th ACM international conference on information and knowledge management*. 2019.
- [9] Škrlić, Blaž, et al. "tax2vec: Constructing interpretable features from taxonomies for short text classification." *Computer Speech & Language* 65 (2021): 101104.
- [10] Wang, Yanshan, et al. "A clinical text classification paradigm using weak supervision and deep representation." *BMC medical informatics and decision making* 19.1 (2019): 1-13.
- [11] Zheng, Jin, and Limin Zheng. "A hybrid bidirectional recurrent convolutional neural network attention-based model for text classification." *IEEE Access* 7 (2019): 106673-106685.
- [12] Pal, Ankit, Muru Selvakumar, and Malaikannan Sankarasubbu. "Multi-label text classification using attention-based graph neural network." *arXiv preprint arXiv:2003.11644* (2020).]
- [13] Oleynik, Michel, et al. "Evaluating shallow and deep learning strategies for the 2018 n2c2 shared task on clinical text classification." *Journal of the American Medical Informatics Association* 26.11 (2019): 1247-1254.
- [14] Banerjee, Imon, et al. "Comparative effectiveness of convolutional neural network (CNN) and recurrent neural network (RNN) architectures for radiology text report classification." *Artificial intelligence in medicine* 97 (2019): 79-88.

- 
- [15] Pappas, Nikolaos, and James Henderson. "Gile: A generalized input-label embedding for text classification." *Transactions of the Association for Computational Linguistics* 7 (2019): 139-155.
- [16] Du, Jingcheng, et al. "ML-Net: multi-label classification of biomedical texts with deep neural networks." *Journal of the American Medical Informatics Association* 26.11 (2019): 1279-1285.
- [17] Chalkidis, Ilias, et al. "An empirical study on large-scale multi-label text classification including few and zero-shot labels." *arXiv preprint arXiv:2010.01653* (2020).
- [18] Chalkidis, Ilias, et al. "Extreme multi-label legal text classification: A case study in EU legislation." *arXiv preprint arXiv:1905.10892* (2019).
- [19] Zhang, Jingqing, Piyawat Lertvittayakumjorn, and Yike Guo. "Integrating semantic knowledge to tackle zero-shot text classification." *arXiv preprint arXiv:1903.12626* (2019).
- [20] Vinayakumar, R., et al. "DBD: deep learning DGA-based botnet detection." *Deep learning applications for cyber security*. Springer, Cham, 2019. 127-149.
- [21] Mujtaba, Ghulam, et al. "Clinical text classification research trends: systematic literature review and open issues." *Expert systems with applications* 116 (2019): 494-520.
- [22] Arbabi, Aryan, et al. "Identifying clinical terms in medical text using ontology-guided machine learning." *JMIR medical informatics* 7.2 (2019): e12596.
- [23] Chantar, Hamouda, et al. "Feature selection using binary grey wolf optimizer with elite-based crossover for Arabic text classification." *Neural Computing and Applications* 32.16 (2020): 12201-12220.
- [24] Chen, Jiaao, Zichao Yang, and Diyi Yang. "Mixtext: Linguistically-informed interpolation of hidden space for semi-supervised text classification." *arXiv preprint arXiv:2004.12239* (2020).
- [25] Abdar, Moloud, et al. "A review of uncertainty quantification in deep learning: Techniques, applications and challenges." *Information Fusion* 76 (2021): 243-297.
- [26] Luketina, Jelena, et al. "A survey of reinforcement learning informed by natural language." *arXiv preprint arXiv:1906.03926* (2019).
- [27] Zhong, Botao, et al. "Hazard analysis: A deep learning and text mining framework for accident prevention." *Advanced Engineering Informatics* 46 (2020): 101152.
- [28] Niu, Zhaoyang, Guoqiang Zhong, and Hui Yu. "A review on the attention mechanism of deep learning." *Neurocomputing* 452 (2021): 48-62.
- [29] Ni, Jinjie, et al. "Recent advances in deep learning based dialogue systems: A systematic survey." *Artificial intelligence review* (2022): 1-101.
- [30] Charbuty, Bahzad, and Adnan Abdulazeez. "Classification based on decision tree algorithm for machine learning." *Journal of Applied Science and Technology Trends* 2.01 (2021): 20-28.