LEGAL JUDGMENT PREDICTION: A NEURAL NETWORK CLASSIFICATION APPROACH

Pavol Travnik
Unicorn University
Prague, Czech Republic
pavol.travnik@pm.me

June 11, 2024

ABSTRACT

Many researchers have attempted to predict court rulings, but upon closer examination, it becomes evident that these efforts often involve merely classifying existing cases into categories based on metadata. This approach lacks true predictive value as it does not leverage the intricate details within the textual data of court decisions. In this paper, I propose using embeddings for data classification, which serves as an optimal tool for handling complex text data through natural language processing (NLP) techniques. By applying embeddings, we achieve a balanced accuracy of 99% in the classification of court decisions. Furthermore, I suggest using embeddings for predicting the outcomes of appeals based solely on the decisions of district courts, achieving a balanced accuracy of 51% for 3 categories (baseline is 33.3%). This methodology not only enhances the predictive capabilities but also provides deeper insights into judicial decision-making processes.

Keywords Legal Judgment Prediction · Neural Network Classification

1 Introduction

1.1 Background

One of the key outputs of the judicial system in the Slovak Republic is the anonymized court rulings, which are made available to the public in PDF format. These documents contain extensive textual information about various cases, providing transparency and accountability in the legal process. The official repository for these documents is accessible through the Ministry of Justice's website at `https://www.justice.gov.sk/sudy-a-rozhodnutia/sudy/rozhodnutia/`.

Each court ruling document is meticulously categorized and enriched with metadata, offering a comprehensive snapshot of the judicial decision. This structured data is crucial for organizing and accessing the rulings efficiently, aiding in legal research, and facilitating various forms of analysis.

1.2 Problem Statement

The primary aim of this research is to conduct a comprehensive statistical analysis of an existing dataset of court case decisions to determine whether it is possible to classify and predict outcomes based on embeddings created from Slovak language texts. This task is particularly challenging due to the difficulty of training on a minor language like Slovak, which lacks the extensive resources available for more widely spoken languages.

Specifically, this research seeks to address the following questions:

- Classification Feasibility: Can embeddings created from Slovak language texts be effectively used for the classification of court case decisions? This involves evaluating whether these embeddings can capture the necessary nuances and complexities of legal texts to achieve accurate classification.

- Predictive Analysis: Is it possible to predict the outcomes of appeals based solely on the decisions of district courts using these embeddings? This aspect of the research aims to develop predictive models that can forecast appeal resolutions with a high degree of accuracy.

1.3 Objectives

1. Dataset Analysis: Conduct a thorough statistical analysis of the existing dataset of Slovak court case decisions to understand its structure, distribution, and key characteristics. This will include an examination of the metadata and textual content to identify patterns and trends.
2. Embedding Creation: Develop and fine-tune embeddings from the Slovak language texts of court decisions. This involves leveraging advanced natural language processing (NLP) techniques to create meaningful representations of the textual data that can capture the nuances and complexities of legal language.
3. Classification Performance: Evaluate the feasibility of using these embeddings for the classification of court case decisions. This will involve training neural network models on a randomized sample of the dataset and measuring relevant performance metrics, such as accuracy, precision, recall, F1 score, and balanced accuracy, to determine the effectiveness of the embeddings in classifying legal texts.
4. Predictive Modeling: Investigate the potential of using embeddings to predict the outcomes of appeals based solely on the decisions of district courts. This will include the development and validation of predictive models, with performance measured using metrics such as balanced accuracy, to assess the models' ability to forecast appeal resolutions accurately.
5. Comparative Analysis: Perform a comparative analysis of the neural network models against baseline models to highlight the improvements achieved through the use of embeddings. This will help in understanding the added value of NLP techniques in the classification and prediction tasks.

1.4 Contributions

If successful, the ability to predict the outcomes of appeals could provide significant insights into the judicial process. It would enable stakeholders to assess whether pursuing an appeal is likely to be beneficial or if accepting the decision of the district court would be a more prudent and cost-effective option. This could lead to substantial savings in both private and public administrative costs, optimizing resource allocation within the judicial system.

2 Literature Review

2.1 Overview of Related Work

Legal Judgment Prediction (LJP) aims to predict court decisions using natural language processing (NLP) and machine learning (ML). Several notable studies have attempted to classify and predict court rulings using various methodologies. The most important phase is source of data. Indian Legal Document Corpus (ILDC) (Malik et al. [2021]) aimed to create a dataset for prediction task. Some studies used ILDC as training set. (Alghazzawi et al. [2022]) Two prominent examples include "Predicting the Law Area and Decisions of French Supreme Court Cases" (Şulea et al. [2017]) and "Swiss-Judgment-Prediction: A Multilingual Legal Judgment Prediction Benchmark" (Niklaus et al. [2021]).

2.2 Gaps in Existing Research

Despite the high accuracy reported in these studies, they have faced significant criticism for not fulfilling the promise of true predictive modeling. As highlighted in "Legal Judgment Prediction: If You Are Going to Do It, Do It Right" (Medvedeva and Mcbride [2023]), many of these legal judgment prediction studies have not accurately predicted court rulings. Instead, they have primarily focused on classifying existing cases based on metadata, which is a fundamentally different task or attempt to make prediction from reasoning of the court in the judgement which can be heavily biased and carry information which party will be successful in dispute. The paper emphasizes the need for using appropriate data, explainability, and application-centric approaches. Many studies use the text of judgments, which are not available until after the decision, making the data inappropriate for prediction. For correct prediction must be avoided reliance on facts extracted from final judgments.

To avoid methodological pitfalls, autor suggest to use for prediction only data prior to the decision, such as the text of the claim, the text of the response or district court decision. This approach is more challenging, as it requires predicting the outcome based on the arguments of the parties or court ruling prior to appeal or supreme court ruling. This approach is more appropriate for true predictive modeling and actually can bring real value to the legal system.

3 Data

3.1 Data aquisition

The data for this research is collected from publicly available court rulings via the API provided by the Ministry of Justice of the Slovak Republic. The specific API endpoint used for downloading the data is `https://obcan.justice.sk/pilot/api/ress-isu-service/v1/rozhodnutie`. This API allows access to a comprehensive collection of anonymized court decisions, including extensive textual information in dokument.url and other relevant metadata.

3.1.1 Metadata

Metadata are scraped using python and scrapy library. Example of metadata response capturing the structure and data types in JSON format is shown below:


```
{
    "guid": "4dbdb3e9-24fa-4762-a37f-f1bf27cb035a:37b90148-844b-48ac-a2b2-784152da1a59",
    "formaRozhodnutia": "Rozsudok",
    "povaha": [
        "Prvostupňové nenapadnuté opravnými prostriedkami"
    ],
    "sud": {
        "registreGuid": "sud_102",
        "nazov": "Mestský súd Bratislava I"
    },
    "sudca": {
        "registreGuid": "sudca_null",
        "meno": "JUDr. Tomáš Škultéty"
    },
    "identifikacneCislo": "1124011089",
    "spisovaZnacka": "14T/9/2024",
    "datumVydania": "05.06.2024",
    "ecli": "ECLI:SK:MSBA1:2024:1124011089.3",
    "oblast": [
        "Trestné právo"
    ],
    "podOblast": [
        "Majetok"
    ],
    "odkazovanePredpisy": [
        {
            "nazov": "/SK/ZZ/2015/300/#paragraf-219.odsek-2",
            "url": "https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2015/300/#paragraf-219.odsek-2"
        }
    ],
    "dokument": {
        "name": "Rozsudok_14T-9-2024.pdf",
        "fileExtension": "PDF",
        "size": 65665,
        "url": "https://obcan.justice.sk/content/public/item/37b90148-844b-48ac-a2b2-784152da1a59"
    },
    "updateDate": "06.06.2024"
}
```


Figure 1: example of JSON metadata response

All collected data is stored in a MongoDB database, preserving the same JSON NoSQL format as received from the API. Snapshot of public dataset contains 3.879.798 items. This approach is chosen to:

- **Maintain Data Integrity**: Ensuring that the original structure and content of the data are retained, preventing potential data flaws during the preprocessing stages.
- **Facilitate Future Selection and Aggregation**: The flexible schema of MongoDB allows for efficient querying, selection, and aggregation of data, which is essential for subsequent analysis and model training.

3.2 Exploratory Data Analysis

The initial data collection process yielded a total of 3.9 million metadata records of court documents. These records contain various pieces of information essential for our analysis and model training. The metadata fields include:

Metadata contains relevant following fields:

Name	Field Name	Description
Court	sud	The court where the decision was made.
Judge	sudca	The judge who issued the ruling.
Unique File Identifier	identifikacneCislo	A unique identifier grouping all documents related to a single case.
Unique Court Identifier	spisovaZnacka	A unique identifier for the court.
Date of Judgment	datumVydania	The date when the judgment was issued.
Form of Decision	formaRozhodnutia	Procedural information about the form of the decision.
Final Decision	povaha	The nature of the final decision.

3.3 Data filtering

The structure of courts in the dataset can be identified by specific strings in the court names:

- **District Courts**: Identified by the string "Okresný súd*".
- **Courts of Appeal**: Identified by the string "Krajský súd*".

By aggregating all unique file identifiers, we can categorize the cases into three main groups:

- Simple Cases (< 2): There is only one file with unique file identifier in dataset.
- **Relevant Cases** (= 2): There are exactly two files with unique file identifier in dataset.
- Complicated Cases (> 2): There are more then two files with unique file identifier in dataset.

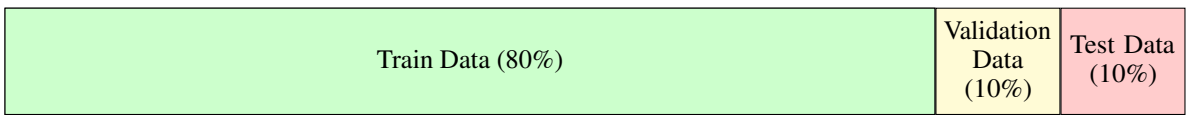
For our prediction model, we specifically need cases that meet the following criteria:

- **Exactly Two Unique File Identifiers**: These cases involve an appeal and have two related decisions, one from the district court and one from the court of appeal.
- **Chronological Order**: The date of the district court's judgment must precede the date of the appeal court's judgment.
- **Final Decision**: The final decision of the district court must be "Prvostupňové nenapadnuté opravnými prostriedkami" (first-instance decision not challenged by remedies).

3.4 Feature Selection

3.5 Data Augmentation

- Final dataset was made by aggregation based on pattern described in appendix.
- Dataset was **not chunked**, due to the size of the dataset it wasn't necessary. For bigger datasets it is recommended to chunk the data to smaller parts to avoid memory issues.
- Dataset was **shuffled** to avoid any potential biases in the data.
- Dataset is **inbalanced** and this has to be handeled in future steps.
- Dataset is splitted to train, val and test sets in ratio **80:10:10**. Train set is solely for training purposes, validation set is used for hyperparameter tuning and test set is used for final evaluation of the model. Final model has no information about test data and serves as a final benchmark for evaluation of the model.



After aggregation made 7 on all scraped data, conditions were applied for filtering only relevant cases. This dataset was made public for future research and is available on `https://huggingface.co/datasets/pavoltravnik/obcan_justice_sk`.

3.5.1 Embeddings

For converting string text for each file were used 4 embeddings models. It was taken into account Massive Text Embedding Benchmark (MTEB) Leaderboard (Muennighoff et al. [2022]) which provides a comprehensive overview of the performance of various embeddings models on a wide range of tasks `https://huggingface.co/spaces/mteb/leaderboard`.

Model	Rank	MaxInput	OutputDim
sentence-transformers/all-MiniLM-L12-v2	99	512	384
avsolatorio/GIST-small-Embedding-v0	36	512	384
openai/text-embedding-3-small	42	8191	1536
openai/text-embedding-3-large	17	8191	3072

3.5.2 Dataset structure

Final data structure of the dataset is shown below:

Model	Rank	MaxInput	OutputDim
sentence-transformers/all-MiniLM-L12-v2	99	512	384
avsolatorio/GIST-small-Embedding-v0	36	512	384
openai/text-embedding-3-small	42	8191	1536
openai/text-embedding-3-large	17	8191	3072

4 Model Architecture

For this research was used Supervised machine learning with feed forward neural newtwork.

4.1 Randomized hyperparameter search

For randomized hyperparameter search was used Optuna finding best combination of hyperparameters for final model training. Based on Optuna trials was chosen this model architecture:

- Input Layer with # of units based on OutputDim of embeddings with ReLU activation
- Output Layer with # of classes
- Softmax activation (Subramanian [2018])

It is simple feed forward neural network with none hidden or dropout layer.

5 Results

5.1 Ethical Considerations

The use of LJP systems must be evaluated for potential benefits and harms. Systems should not undermine the role of human judgment in the legal process but serve just as another tool for better decision making and examination. It is vital to get more information from transparent and public legal corpuses via neural networks and they can serve as state of art diagnostic tool for legal professionals to examine patterns, similarities and potential flaws in jurisprudence.

5.2 Performance Analysis of Classification

Classification is based on information which is avalable prior the decision and is made for fields [district law field], [district decision form], [district sud]. The model can capture the data and classify it with high accuracy. This is caused, that model easily can capture the patterns in the data and then classify them based on existing information easily.

Performance Analysis of Classification achieves balanced accuracy near 99% just because data contains patterns which are easily classifiable.

The results are presented in the following tables and figures.

5.2.1 Performance Analysis of label [district law field]

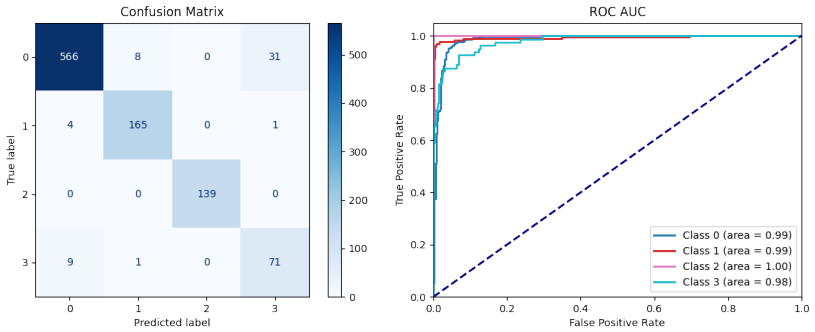


Figure 2: Model: openai/text-embedding-3-large

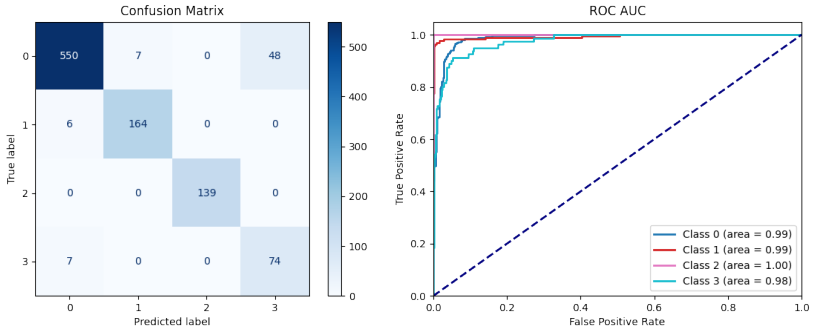


Figure 3: Model: openai/text-embedding-3-small

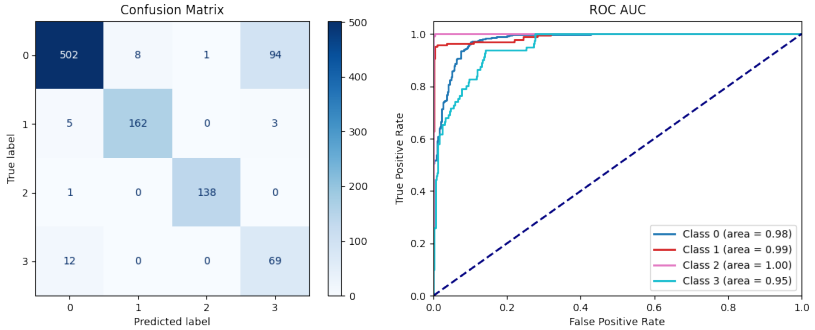


Figure 4: Model: avsolatorio/GIST-small-Embedding-v0

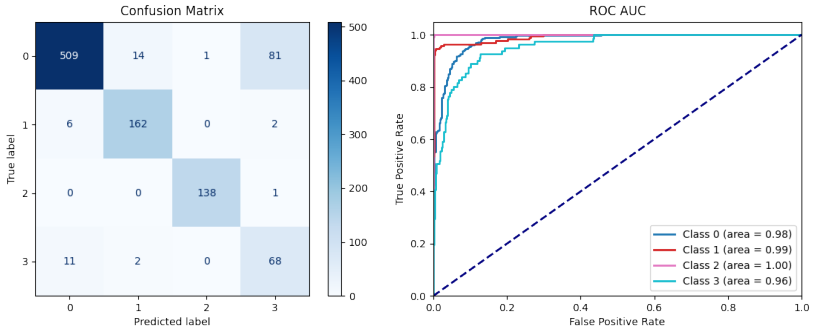


Figure 5: Model: sentence-transformers/all-MiniLM-L12-v2

Model	Task	Accuracy	Bal. Acc	F1 Score	Precision	Recall
openai/text-embedding-3-large	classification	0.95	0.95	0.95	0.95	0.95
openai/text-embedding-3-small	classification	0.93	0.95	0.94	0.95	0.93
sentence-transformers/all-MiniLM-L12-v2	classification	0.88	0.91	0.89	0.92	0.88
avsolatorio/GIST-small-Embedding-v0	classification	0.88	0.91	0.89	0.92	0.88

5.2.2 Performance Analysis of label [district decision form]

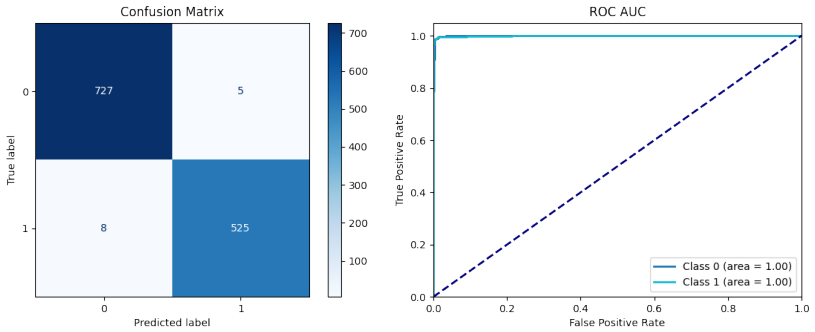


Figure 6: Model: openai/text-embedding-3-large

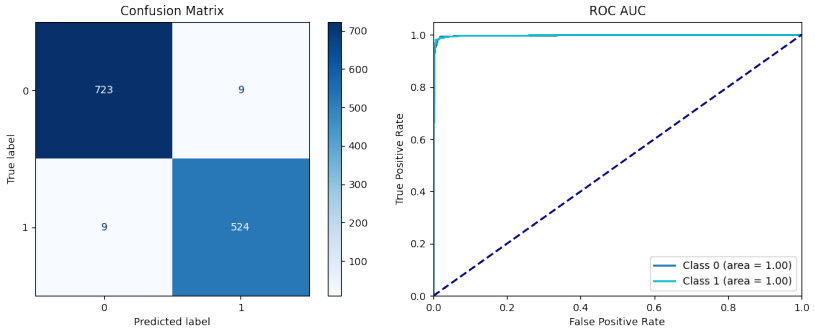


Figure 7: Model: openai/text-embedding-3-small

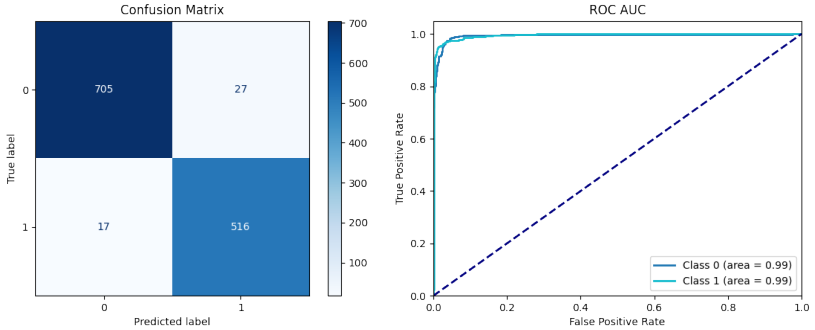


Figure 8: Model: avsolatorio/GIST-small-Embedding-v0

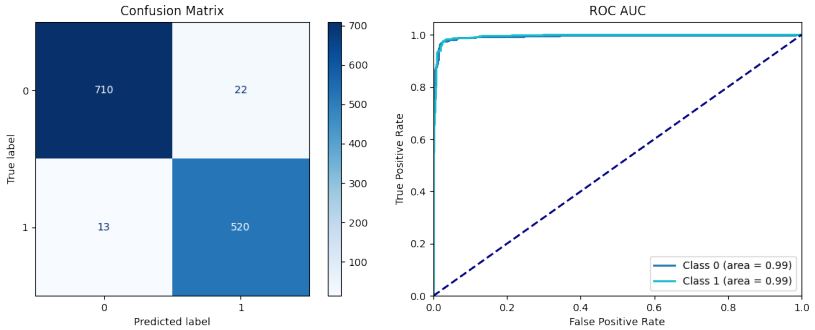


Figure 9: Model: sentence-transformers/all-MiniLM-L12-v2

Model	Task	Accuracy	Bal. Acc	F1 Score	Precision	Recall
openai/text-embedding-3-large	classification	0.99	0.99	0.99	0.99	0.99
openai/text-embedding-3-small	classification	0.99	0.99	0.99	0.99	0.99
sentence-transformers/all-MiniLM-L12-v2	classification	0.97	0.97	0.97	0.97	0.97
avsolatorio/GIST-small-Embedding-v0	classification	0.97	0.97	0.97	0.97	0.97

5.2.3 Performance Analysis of label [district sud]

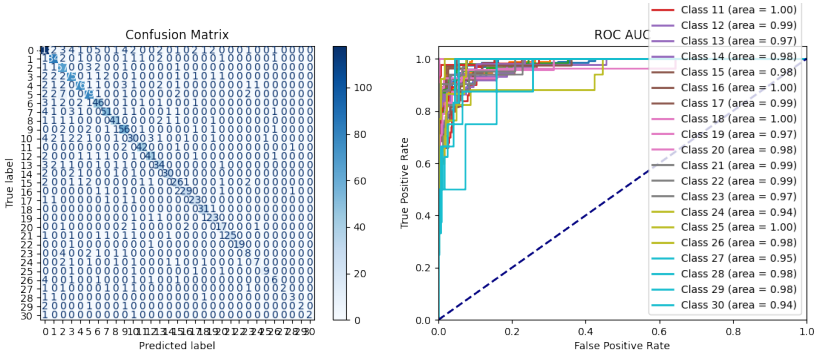


Figure 10: Model: openai/text-embedding-3-large

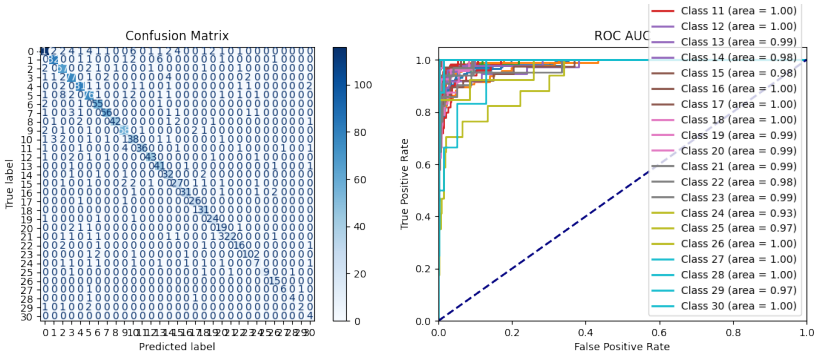


Figure 11: Model: openai/text-embedding-3-small

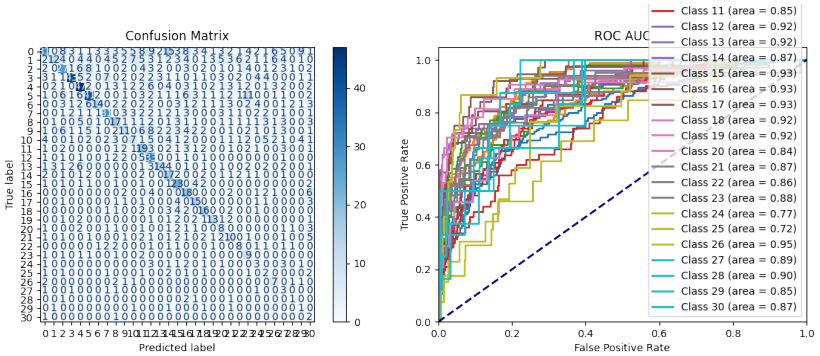


Figure 12: Model: avsolatorio/GIST-small-Embedding-v0

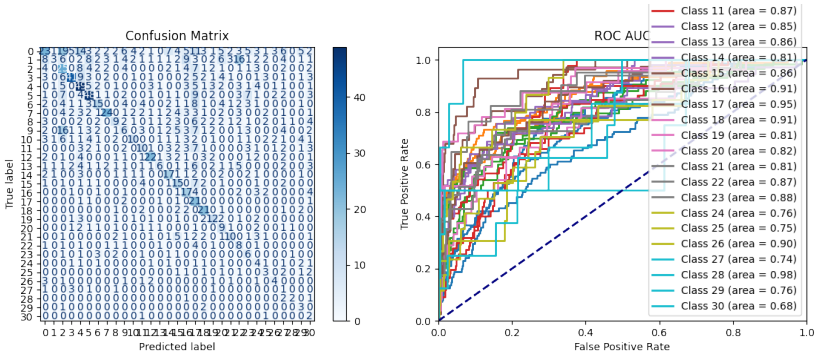


Figure 13: Model: sentence-transformers/all-MiniLM-L12-v2

Model	Task	Accuracy	Bal. Acc	F1 Score	Precision	Recall
openai/text-embedding-3-large	classification	0.8	0.73	0.8	0.8	0.8
openai/text-embedding-3-small	classification	0.82	0.79	0.82	0.83	0.82
sentence-transformers/all-MiniLM-L12-v2	classification	0.33	0.36	0.33	0.42	0.33
avsolatorio/GIST-small-Embedding-v0	classification	0.28	0.29	0.27	0.35	0.28

5.3 Performance Analysis of Prediction

Prediction in opposite to classification are based on data which are not known in the time of prediction. The model must rely solely on previous information. The results are presented in the following tables and figures.

In some cases for fields like [appeal sud] or [appeal law field] is very easy to predict metadata of the appeal decision, because are tightly connected to existing procedural rules like for field [appeal sud] (if you know district court, there is very high probability which court will be the appeal court) or for field [appeal law field] for which if you know context of the matter, it is unlikely it changes in appeal.

The true challenge for prediction is however [appeal final decision] field which is the crutial to determine, if the case is going to be held or dismissed.

The results are presented in the following tables and figures.

5.3.1 Performance Analysis of label [appeal sud]

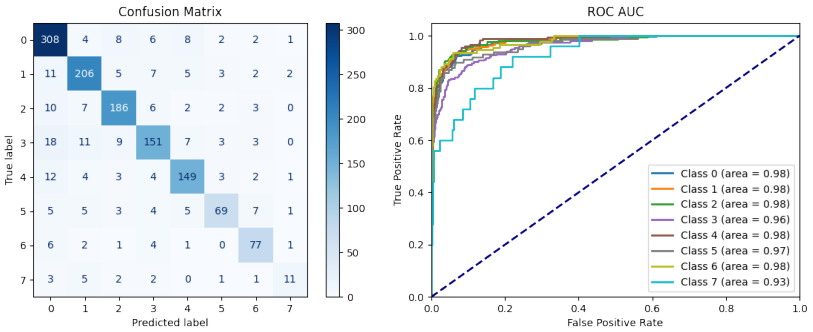


Figure 14: Model: openai/text-embedding-3-large

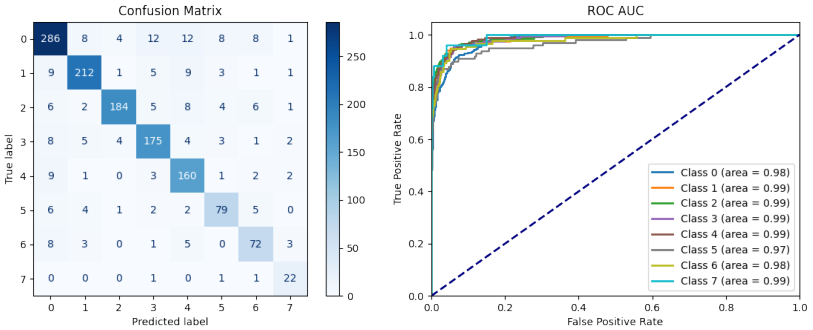


Figure 15: Model: openai/text-embedding-3-small

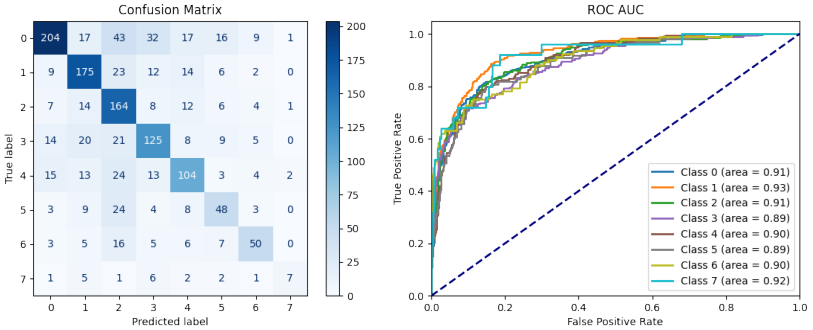


Figure 16: Model: avsolatorio/GIST-small-Embedding-v0

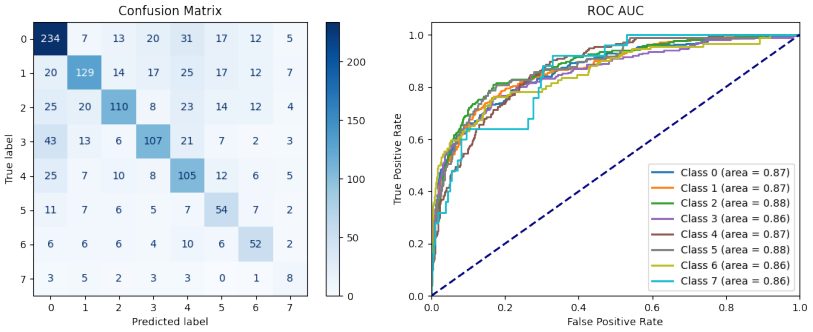


Figure 17: Model: sentence-transformers/all-MiniLM-L12-v2

Model	Task	Accuracy	Bal. Acc	F1 Score	Precision	Recall
openai/text-embedding-3-large	prediction	0.83	0.77	0.83	0.83	0.83
openai/text-embedding-3-small	prediction	0.85	0.85	0.86	0.86	0.85
sentence-transformers/all-MiniLM-L12-v2	prediction	0.57	0.54	0.58	0.59	0.57
avsolatorio/GIST-small-Embedding-v0	prediction	0.63	0.57	0.63	0.65	0.63

5.3.2 Performance Analysis of label [appeal decision form]

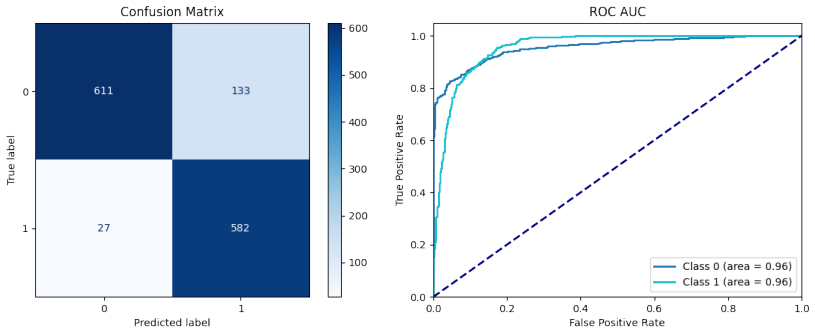


Figure 18: Model: openai/text-embedding-3-large

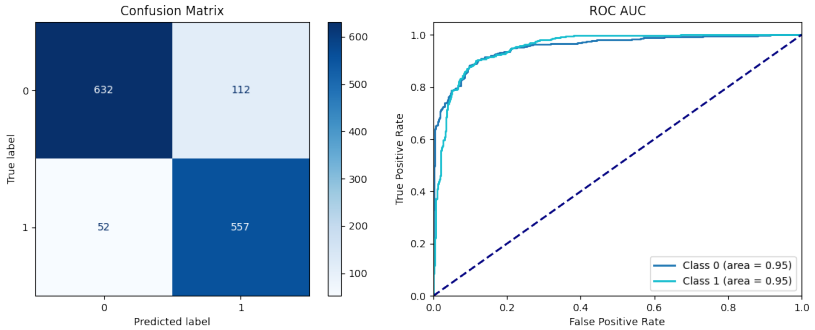


Figure 19: Model: openai/text-embedding-3-small

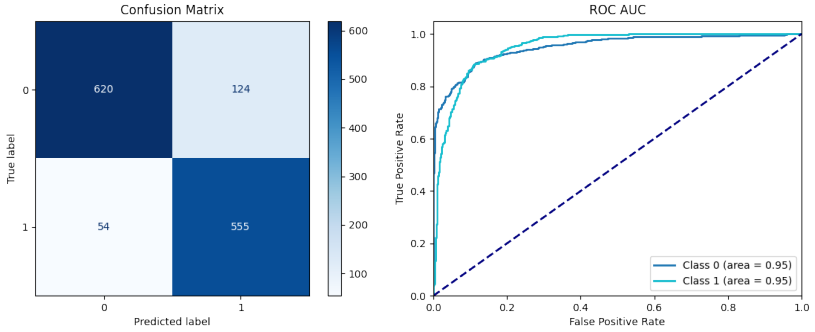


Figure 20: Model: avsolatorio/GIST-small-Embedding-v0

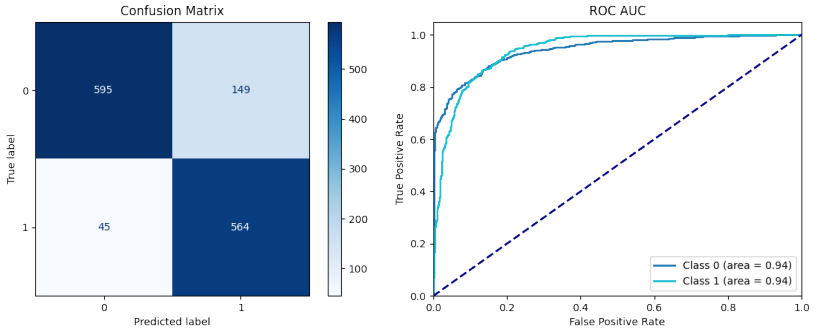


Figure 21: Model: sentence-transformers/all-MiniLM-L12-v2

Model	Task	Accuracy	Bal. Acc	F1 Score	Precision	Recall
openai/text-embedding-3-large	prediction	0.88	0.89	0.88	0.89	0.88
openai/text-embedding-3-small	prediction	0.88	0.88	0.88	0.88	0.88
sentence-transformers/all-MiniLM-L12-v2	prediction	0.86	0.86	0.86	0.87	0.86
avsolatorio/GIST-small-Embedding-v0	prediction	0.87	0.87	0.87	0.87	0.87

5.3.3 Performance Analysis of label [appeal law field]

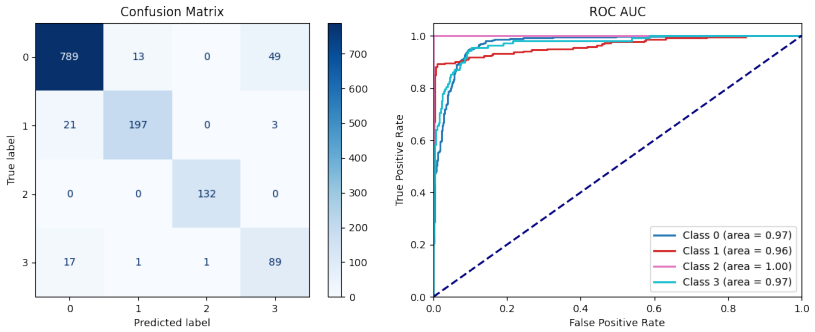


Figure 22: Model: openai/text-embedding-3-large

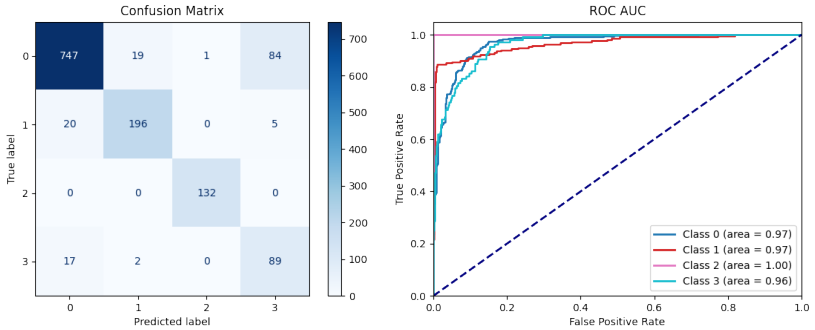


Figure 23: Model: openai/text-embedding-3-small

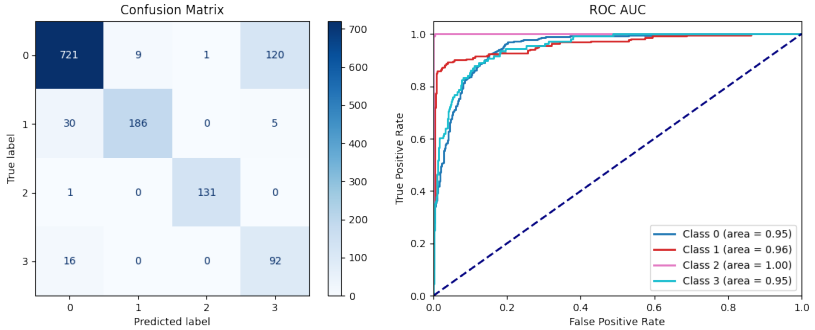


Figure 24: Model: avsolatorio/GIST-small-Embedding-v0

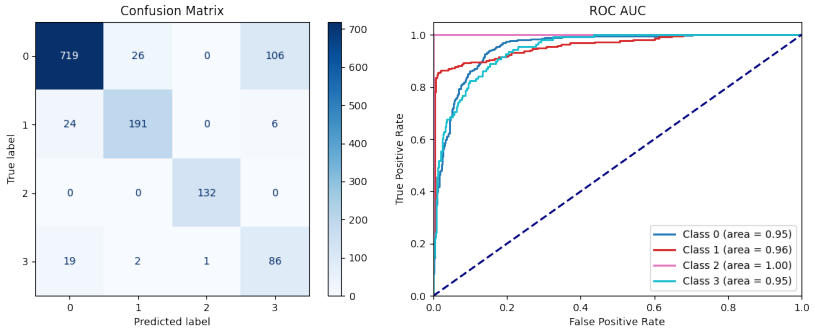


Figure 25: Model: sentence-transformers/all-MiniLM-L12-v2

Model	Task	Accuracy	Bal. Acc	F1 Score	Precision	Recall
openai/text-embedding-3-large	prediction	0.92	0.91	0.92	0.93	0.92
openai/text-embedding-3-small	prediction	0.89	0.9	0.89	0.91	0.89
sentence-transformers/all-MiniLM-L12-v2	prediction	0.86	0.88	0.87	0.89	0.86
avsolatorio/GIST-small-Embedding-v0	prediction	0.87	0.89	0.88	0.91	0.87

5.3.4 Performance Analysis of label [appeal final decision]

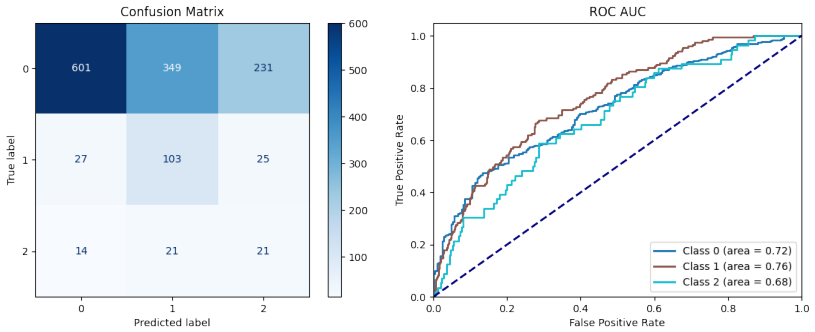


Figure 26: Model: openai/text-embedding-3-large

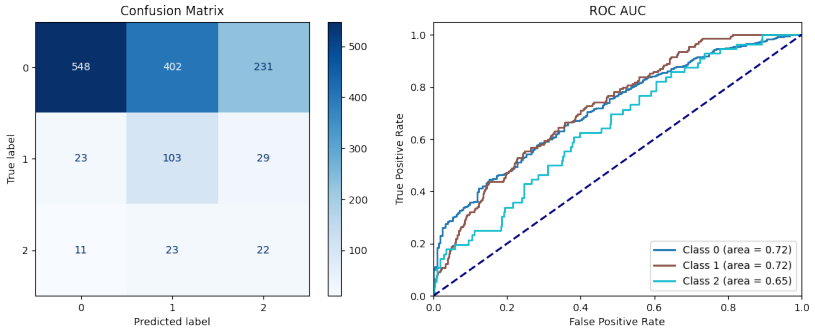


Figure 27: Model: openai/text-embedding-3-small

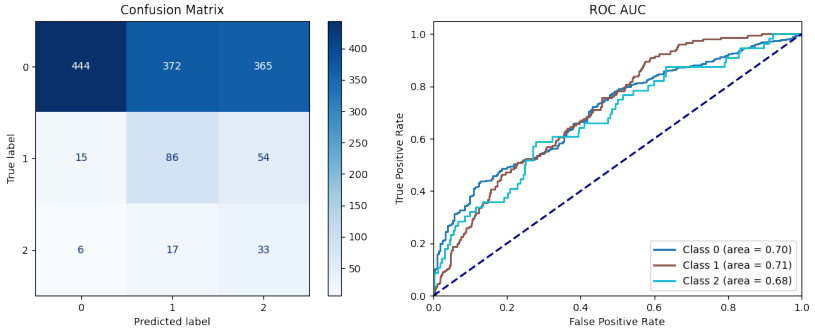


Figure 28: Model: avsolatorio/GIST-small-Embedding-v0

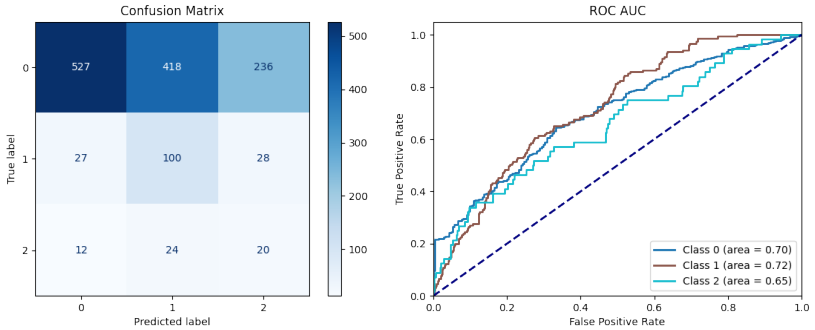


Figure 29: Model: sentence-transformers/all-MiniLM-L12-v2

Model	Task	Accuracy	Bal. Acc	F1 Score	Precision	Recall
openai/text-embedding-3-large	prediction	0.52	0.51	0.6	0.82	0.52
openai/text-embedding-3-small	prediction	0.41	0.47	0.49	0.81	0.41
sentence-transformers/all-MiniLM-L12-v2	prediction	0.45	0.49	0.53	0.82	0.45
avsolatorio/GIST-small-Embedding-v0	prediction	0.43	0.51	0.52	0.82	0.43

5.4 Comparative Analysis

5.5 Performance Analysis of Pediction on randomized labels

Comparative analysis is held solely agains randomized data in which input data stays same and labels are shuffled. This helps us truly understand the teoretical baseline of the model.

5.5.1 Performance Analysis of Pediction on randomized labels [appeal final decision shuffled]

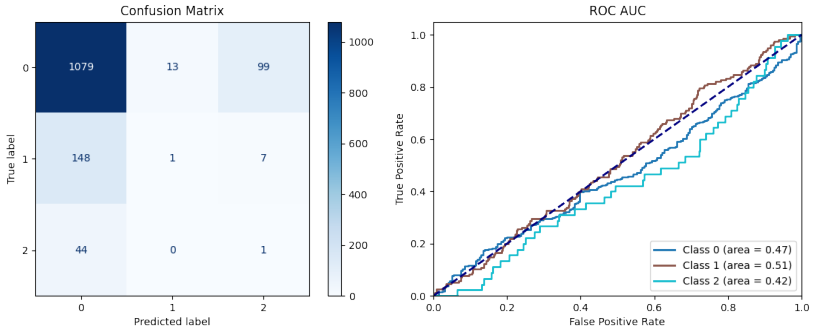


Figure 30: Model: openai/text-embedding-3-large

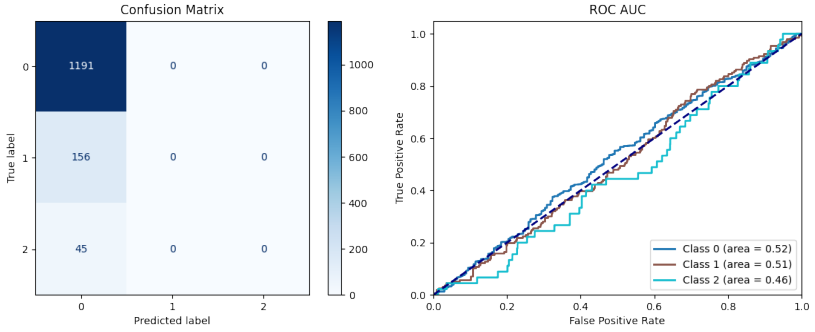


Figure 31: Model: openai/text-embedding-3-small

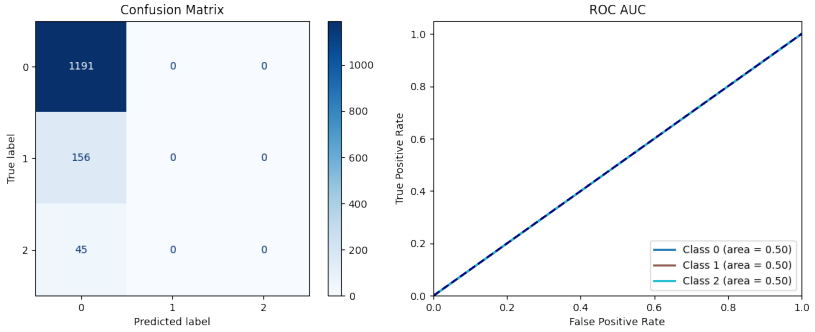


Figure 32: Model: avsolatorio/GIST-small-Embedding-v0

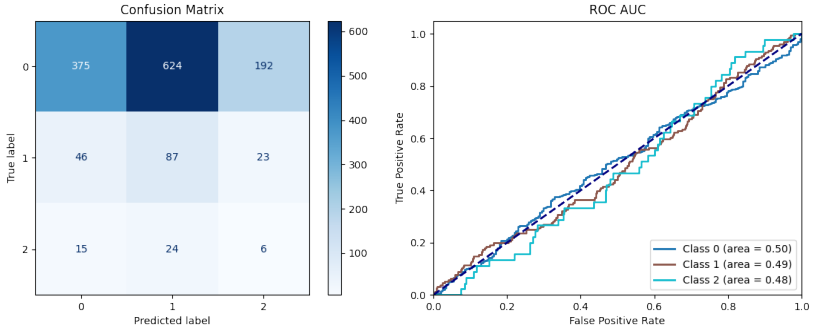


Figure 33: Model: sentence-transformers/all-MiniLM-L12-v2

Model	Task	Accuracy	Bal. Acc	F1 Score	Precision	Recall
openai/text-embedding-3-large	prediction	0.78	0.31	0.75	0.73	0.78
openai/text-embedding-3-small	prediction	0.86	0.33	0.79	0.73	0.86
sentence-transformers/all-MiniLM-L12-v2	prediction	0.86	0.33	0.79	0.73	0.86
avsolatorio/GIST-small-Embedding-v0	prediction	0.86	0.33	0.79	0.73	0.86

6 Conclusion

This study declares, that for NLP task as Legal Judgment Prediction is necessary to distinquish between classification, which is based on data which can easily classify court decisions and prediction, which rely solely on previous information and future information are not known. Classification usually get very high balanced accuracy for inbalanced datasets, however prediction is much more difficult task. This study shows that embeddings can be used for classification with extremly high accuracy, but for prediction the results are not as good as expected, but even though they support hypothesis, that appeal can be with some degree predicted solely on district court ruling. As embeddings are changing text to numerical representation in high dimensional space, the information why this model is able to predict some rulings are lost. This training was made just on very small dataset and contains only public data. The capability of the model can hugely increase if model has all available data about the case like fillings of each party, metadata for each case file are correct and available and connection between cases are known. There were used multiple embedding models, as can be seen on data, for simple classification with simple patterns is actually better to use existing opensource models with small dimensions. For prediction and to capture meaning in the data is better to use bigger models trained also on the target language.

This study supports, that embeddings with connection to custom NN have some predictive power and autor suggest not only to make more research in this field, but also provide empirical evidence, that legal judgment prediction based on public dataset is feasible. For this purpose more transparent justice system is needed, which will provide more data for training such models.

Author needs to emphasize, that this study can significantly improve by bigger dataset, abanoning mean pooling, getting more metadata and using more advanced model architectures.

7 Appendix

Aggregation script in MongoDB to produce the dataset for the analysis:

```
[
    {
      $group: {
        // Step 1: Group by identifikacneCislo
        _id: "$item.identifikacneCislo",
        items: {
          $push: "$$ROOT",
        },
        count: {
          $sum: 1,
        },
      },
    },
    {
      $match: {
        count: 2,
      }, // Step 2: Filter groups with exactly two items
    },
    {
      $match: {
        // Step 3: Ensure one item has sud.nazov starting with "Krajský súd" and the other with "Okresný súd"
        "items.0.item.sud.nazov": {
          $regex: "^Krajský súd",
        },
        "items.1.item.sud.nazov": {
          $regex: "^Okresný súd",
        },
      },
    },
    {
      $project: {
        identifikacneCislo: "$_id",
        district_final_decision: {
          $arrayElemAt: [
            {
```

```
      $arrayElemAt: [
        "$items.item.povaha",
        1,
      ],
    },
    0,
  ],
},
district_law_field: {
  $arrayElemAt: [
    {
      $arrayElemAt: [
        "$items.item.oblast",
        1,
      ],
    },
    0,
  ],
},
district_decision_form: {
  $arrayElemAt: [
    "$items.item.formaRozhodnutia",
    1,
  ],
},
district_spzn: {
  $arrayElemAt: [
    "$items.item.spisovaZnacka",
    1,
  ],
},
district_judge: {
  $arrayElemAt: [
    "$items.item.sudca.meno",
    1,
  ],
},
district_judge_id: {
  $arrayElemAt: [
    "$items.item.sudca.registreGuid",
    1,
  ],
},
district_sud: {
  $arrayElemAt: [
    "$items.item.sud.nazov",
    1,
  ],
},
district_guid: {
  $arrayElemAt: ["$items.item.guid", 1],
},
district_date_judgement: {
  $arrayElemAt: [
    "$items.item.datumVydania",
    1,
  ],
},
district_dokumenturl: {
  $arrayElemAt: [
    "$items.item.dokument.url",
    1,
  ],
},
appeal_final_decision: {
```

```
      $arrayElemAt: [
        {
          $arrayElemAt: [
            "$items.item.povaha",
            0,
          ],
        },
        0,
      ],
    },
    appeal_law_field: {
      $arrayElemAt: [
        {
          $arrayElemAt: [
            "$items.item.oblast",
            0,
          ],
        },
        0,
      ],
    },
    appeal_decision_form: {
      $arrayElemAt: [
        "$items.item.formaRozhodnutia",
        0,
      ],
    },
    appeal_spzn: {
      $arrayElemAt: [
        "$items.item.spisovaZnacka",
        0,
      ],
    },
    appeal_judge: {
      $arrayElemAt: [
        "$items.item.sudca.meno",
        0,
      ],
    },
    appeal_judge_id: {
      $arrayElemAt: [
        "$items.item.sudca.registreGuid",
        0,
      ],
    },
    appeal_sud: {
      $arrayElemAt: [
        "$items.item.sud.nazov",
        0,
      ],
    },
    appeal_guid: {
      $arrayElemAt: ["$items.item.guid", 0],
    },
    appeal_date_judgement: {
      $arrayElemAt: [
        "$items.item.datumVydania",
        0,
      ],
    },
    appeal_dokumenturl: {
      $arrayElemAt: [
        "$items.item.dokument.url",
        0,
      ],
```

```
        },
        _id: 0,
      },
    },
    {
      $addFields: {
        district_date_judgement: {
          $dateFromString: {
            dateString: "$district_date_judgement",
            format: "%d.%m.%Y",
          },
        },
        appeal_date_judgement: {
          $dateFromString: {
            dateString: "$appeal_date_judgement",
            format: "%d.%m.%Y",
          },
        },
      },
    },
    {
      $match: {
        $expr: {
          $lt: [
            "$district_date_judgement",
            "$appeal_date_judgement",
          ],
        },
      },
    },
  ]
```

References

Vijit Malik, Rishabh Sanjay, Shubham Kumar Nigam, Kripabandhu Ghosh, Shouvik Kumar Guha, Arnab Bhattacharya, and Ashutosh Modi. ILDC for CJPE: Indian legal documents corpus for court judgment prediction and explanation. In Chengqing Zong, Fei Xia, Wenjie Li, and Roberto Navigli, editors, *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 4046–4062, Online, August 2021. Association for Computational Linguistics. doi:10.18653/v1/2021.acl-long.313. URL https://aclanthology.org/2021.acl-long.313.

Daniyal Alghazzawi, Omaimah Bamasag, Aiiad Albeshri, Iqra Sana, Hayat Ullah, and Muhammad Zubair Asghar. Efficient prediction of court judgments using an lstm+cnn neural network model with an optimal feature set. *Mathematics*, 10(5), 2022. ISSN 2227-7390. doi:10.3390/math10050683. URL https://www.mdpi.com/2227-7390/10/5/683.

Octavia-Maria Şulea, Marcos Zampieri, Mihaela Vela, and Josef van Genabith. Predicting the law area and decisions of French Supreme Court cases. In Ruslan Mitkov and Galia Angelova, editors, *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017*, pages 716–722, Varna, Bulgaria, September 2017. INCOMA Ltd. doi:10.26615/978-954-452-049-6_092. URL https://doi.org/10.26615/978-954-452-049-6_092.

Joel Niklaus, Ilias Chalkidis, and Matthias Stürmer. Swiss-judgment-prediction: A multilingual legal judgment prediction benchmark. In Nikolaos Aletras, Ion Androutsopoulos, Leslie Barrett, Catalina Goanta, and Daniel Preotiuc-Pietro, editors, *Proceedings of the Natural Legal Language Processing Workshop 2021*, pages 19–35, Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics. doi:10.18653/v1/2021.nllp-1.3. URL https://aclanthology.org/2021.nllp-1.3.

Masha Medvedeva and Pauline Mcbride. Legal judgment prediction: If you are going to do it, do it right. In Daniel Preoțiuc-Pietro, Catalina Goanta, Ilias Chalkidis, Leslie Barrett, Gerasimos (Jerry) Spanakis, and Nikolaos Aletras, editors, *Proceedings of the Natural Legal Language Processing Workshop 2023*, pages 73–84, Singapore, December 2023. Association for Computational Linguistics. doi:10.18653/v1/2023.nllp-1.9. URL https://aclanthology.org/2023.nllp-1.9.

Niklas Muennighoff, Nouamane Tazi, Loïc Magne, and Nils Reimers. Mteb: Massive text embedding benchmark. *arXiv preprint arXiv:2210.07316*, 2022. doi:10.48550/ARXIV.2210.07316. URL https://arxiv.org/abs/2210.07316.

V. Subramanian. *Deep Learning with PyTorch: A practical approach to building neural network models using PyTorch*. Packt Publishing, 2018. ISBN 9781788626071. URL https://books.google.cz/books?id=DOlODwAAQBAJ.