Explainable NLP Model for Predicting Patient Admissions at Emergency Department Using Triage Notes

Émilien Arnaud, **Mahmoud Elbattah**, Pedro A. Moreno-Sánchez, Gilles Dequen, Daniel Aiham Ghazali

Université de Picardie Jules Verne (UPJV), France

mahmoud.elbattah@u-picardie.fr

Our University

Université de Picardie Jules Verne, France

Laboratoire Modélisation, Information, Système (MIS Lab)

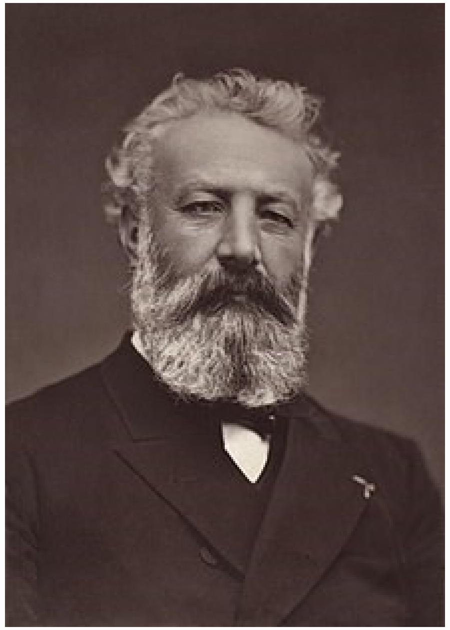
Credits: Google Maps

UNIVERSITÉ de Picardie Jules Verne

About Jules Verne

- Jules Verne (8 February 1828 – 24 March 1905) was a French novelist, poet, and playwright.

- Famous for the Journey to the Center of the Earth.



Jules Verne, French Novelist

Maison de Jules Verne, Amiens, France

Journey to the Center of the Earth, 2008

Sources of Images:
https://en.wikipedia.org/wiki/Jules_Verne
https://fr.wikipedia.org/wiki/Maison_de_Jules_Verne
https://en.wikipedia.org/wiki/Journey_to_the_Center_of_the_Earth_(2008_theatrical_film)

Study Context

- A use case of triage notes in the French language, from the University Hospital of Picardie Jules Verne in France.

- Exploring pre-trained BERT models for learning embeddings from clinical notes for predicting patient hospitalization.

- While also delivering “interpretable” results using the LIME approach.

UNIVERSITÉ de Picardie Jules Verne

Our Earlier Work

- Early prediction of hospitalization[1].
- Prediction of medical specialties for patients hospitalized [2].
- Developing clustering models using BERT-based embeddings [3].

[1] Arnaud, E., Elbattah, M., Gignon, M., & Dequen, G. (2020). Deep learning to predict hospitalization at triage: Integration of structured data and unstructured text. *In Proceedings of the IEEE International Conference on Big Data*. IEEE.

[2] Arnaud, E., Elbattah, M., Gignon, G & Dequen, G. (2021). NLP-based prediction of medical specialties at hospital admission using triage notes. In Proceedings of IEEE International Conference on Healthcare Informatics (ICHI).

[3] Arnaud, E., Elbattah, M., Gignon, G & Dequen, G. (2022). Learning embeddings from free-text triage notes using pretrained transformer models. In Proceedings of the 15th International Joint Conf. on Biomedical Engineering Systems and Technologies (BIOSTEC).

Data Description

- Provided by the University Hospital of Picardie Jules Verne, France.
- About 260K emergency admission records covering 2015 to 2019.
- 9 **categorical** variables: Details on admission (e.g. Origin, Arrival Modality, Family Status, etc.)
- 15 **numeric** variables: Patient's age and other vital signs recorded (e.g. Heart Rate, Temperature, Blood Pressure, etc.).
- 4 **text** variables: Nurse Notes, Psychiatric History, Surgical History, Medical History.
- Binary label: Patient hospitalization.

UNIVERSITÉ de Picardie Jules Verne

Data Description (cont'd)

- We excluded all categorical and numerical variables.
- Only focused on text data.
- For each patient, the 4 text fields were combined to form a single document.
- Procedures of text normalization were applied to clean and standardize the textual data.

UNIVERSITÉ de Picardie Jules Verne

Transformer Models

- **CamemBERT (Martin et al. 2019):** A popular model for the French language developed by the Inria institute. Based on the BERT architecture and has been pretrained on a large French corpus.

- **FlauBERT (Le et al. 2019):** FlauBERT is a variation of the BERT model using a bidirectional attention mechanism to comprehend the context within a sentence. It was specifically built for the French language and has been trained on a corpus of French text.

UNIVERSITÉ de Picardie Jules Verne

Model Fine-Tuning

Model	Params	Embedding Dimension	Runtime
CamemBERT	110M	768	13.3 hrs
FlauBERT	137 M	768	13.5 hrs

Using a bi-GPU node featuring two V-100 Nvidia GPUs.

UNIVERSITÉ de Picardie Jules Verne

Model Performance

Model	AUROC	Precision	Recall	F1-Score
CamemBERT	**0.75**	0.69	**0.70**	**0.70**
FlauBERT	0.74	0.69	0.69	0.69

Model Explainability

- Local Interpretable Model-agnostic Explanations (LIME) (Ribeiro et al. 2016).
- Approximating the model's decision boundary around a specific prediction, and then fitting a simple model to this boundary.
- This simple model can then be used to understand which input features had the most impact on the prediction.
- LIME is **model-agnostic**, meaning it can be used to explain the predictions of any model, regardless of the type of model or algorithm used.

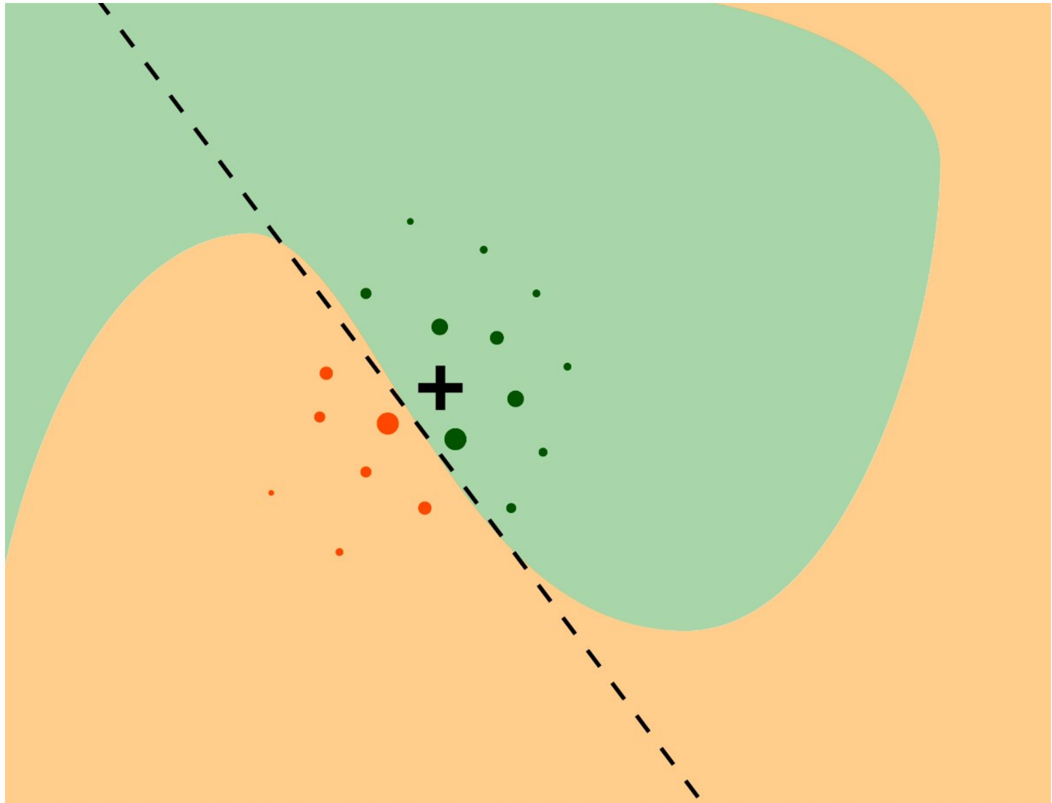


Figure Source: https://ema.drwhy.ai/LIME.html

Ribeiro, Marco Tulio, Sameer Singh, and Carlos Guestrin. "" Why should I trust you?" Explaining the predictions of any classifier." In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135-1144. 2016.

Model Explainability (cont'd)

Effort dyspnea throat constriction no sign of respiratory complication describes palpitations EKG done and checked intermittent palpitations progress since ten days low right chest pain spontaneous resolving recent biology DDM neg.

Figure 1. Example of explaining the prediction of a discharged patient (i.e. True Negative).

UNIVERSITÉ de Picardie Jules Verne

Model Explainability (cont’d)

Adenocarcinoma of the right colic angle operated on February pt r contiguous pancreatitic involvement liver metastases treated by cephalic duopancratectomy right hemicolectomy ileostomy cholecystectomy hepatic wedge segment vii adjuvant treatment with folfox courses mai December weaned smoking since one year oh cirrhosis child B left wrist fracture years amputation fingers iii iv v right hand work accident anorexia since week vomits colicneoplasia evolving w/ meta

Figure 2. Example of explaining the prediction of an admitted patient (i.e. True Positive), probability of admission ≈64%.

UNIVERSITÉ de Picardie Jules Verne

Conclusions

- There were instances where the model explanations were largely aligned with the physician’s understanding.
- However, this was **not** always the case.
- In general ,the physician much appreciated the insights provided by the NLP-based explanations for understanding of the model’s predictions.
- The integration of a human-in-the-loop approach holds promise by leveraging the medical review of explanations.

Thank You!

Please feel free to contact me for any opportunity for collaboration:

mahmoud.elbattah@u-picardie.fr