

NLP-Based Prediction of Medical Specialties at Hospital Admission Using Triage Notes

Émilien Arnaud, Mahmoud Elbattah, Maxime Gignon, Gilles Dequen

Université de Picardie Jules Verne (UPJV), France

mahmoud.elbattah@u-picardie.fr

Overcrowding in Emergency Departments



spitalier Universitaire d'Amiens
gences

Attente	Soins C	Soins L1	Soins L2	SAUPA	UHTCD	SAUV	Zone de transit	Examens	Synthèse	Attente de clôture
Attente IOA			Attente Soins Courts				Attente Soins Longs			
GENEVIEVE 0h 24m MED 56a			HUGO 0h 35m MED 18a				NATAGHA 0h 18m MED 26a			
TANIA 0h 8m INF 36a			PAUL 0h 3m C 34a				ESTHER 0h 20m MED 66a			
JACQUES 0h 2m INF 57a			PAULO 0h 21m MED 41a				SAMUEL 0h 53m MED 55a			
IOA 1 Vide			JEANNE 3h 21m MED 95a				MICHEL 0h 45m MED 18a			
IOA 2 Vide			MAVOME 0h 19m MED 29a				YEDDI 0h 40m MED 24a			
IOA 3 Vide			JEAN 4h 12m MED 87a				MICHAEL 0h 52m MED 44a			
IOA 4 Vide			CLEMEN 1h 45m MED 22a				MAGALI 3h 45m MED 44a			
			SANDRINE 2h 47m MED 42a				NICOLAS 2h 53m MED 45a			

Photo Credits: Dr Émilien Arnaud, CHU, Amiens, France

The Role of Data Analytics

- Early prediction of hospitalization (our earlier work¹).
- Prediction of medical specialties for patients hospitalized.

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.

Data Description

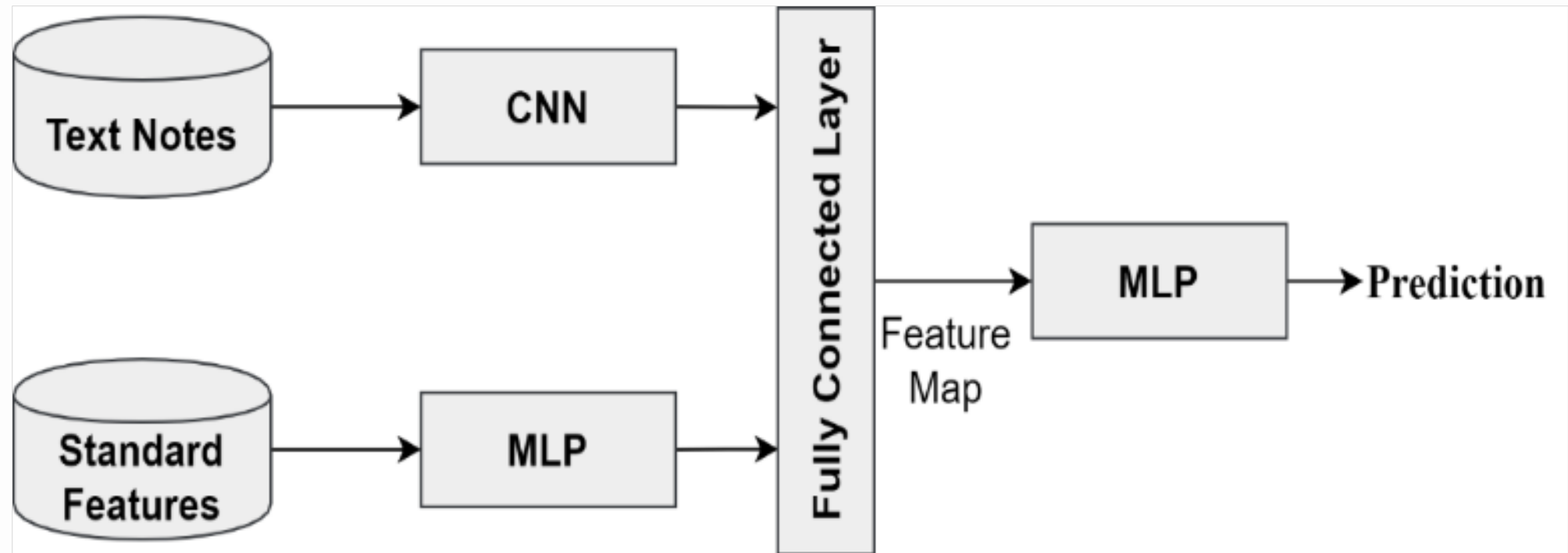
More than 260K ED records over the period of January 2015 to June 2019.

#	Field Name	Type
1	Arrival (Week Day /Hour)	Categorical
2	Gender	Categorical
3	Origin	Categorical
4	Arrival Modlaity	Categorical
5	Accompaniers	Categorical
6	Family Status	Categorical
7	Waiting Modality	Categorical
8	Reason for Encounter	Categorical
9	Circumstances	Categorical
10	Age	Numeric
11	Oxygen Flow	Numeric
12	Heart Rate	Numeric
13	Respiration Rate	Numeric
14	Systolic Blood Pressure	Numeric
15	Diastolic Blood Pressure	Numeric
16	Pain Scale	Numeric
17	Temperature	Numeric
18	Oxygen Saturation	Numeric
19	Capillary Blood Glucose	Numeric
20	Capillary Blood Hemoglobin	Numeric
21	Bladder volume	Numeric
22	Capillary Blood Ketones	Numeric
23	Breath Test of Alcohol	Numeric
24	Nurse Triage Scale	Numeric
25	Nurse Notes	Text
26	Psychiatric History	Text
27	Surgical History	Text
28	Medical History	Text

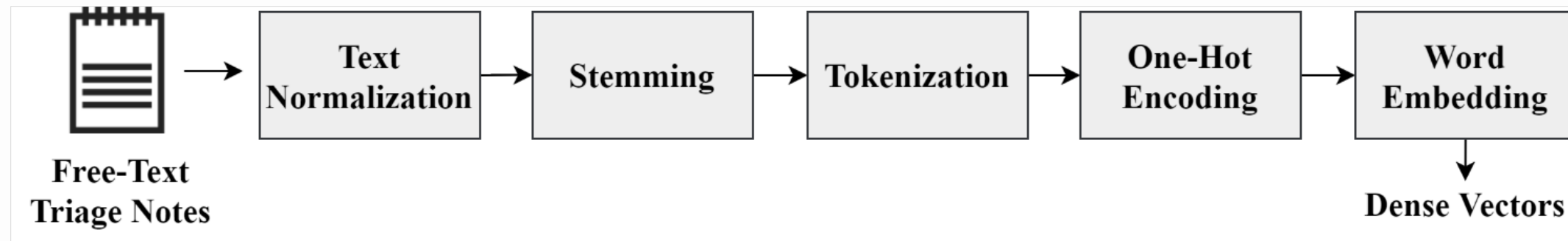
Data Description (cont'd)

Specialty / Label	Hospitalization %
Surgery / CHIR	19.7%
Short-Term Hospitalization Unit / UHCD	42.4%
Medical Specialty / MED	33%
Other	4.9%

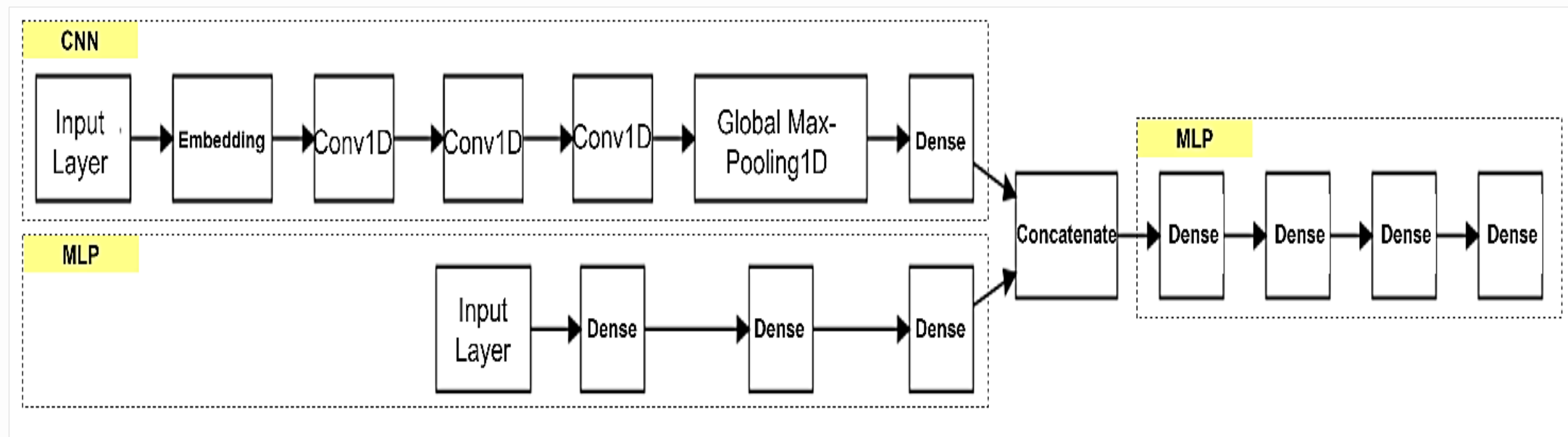
Approach Overview



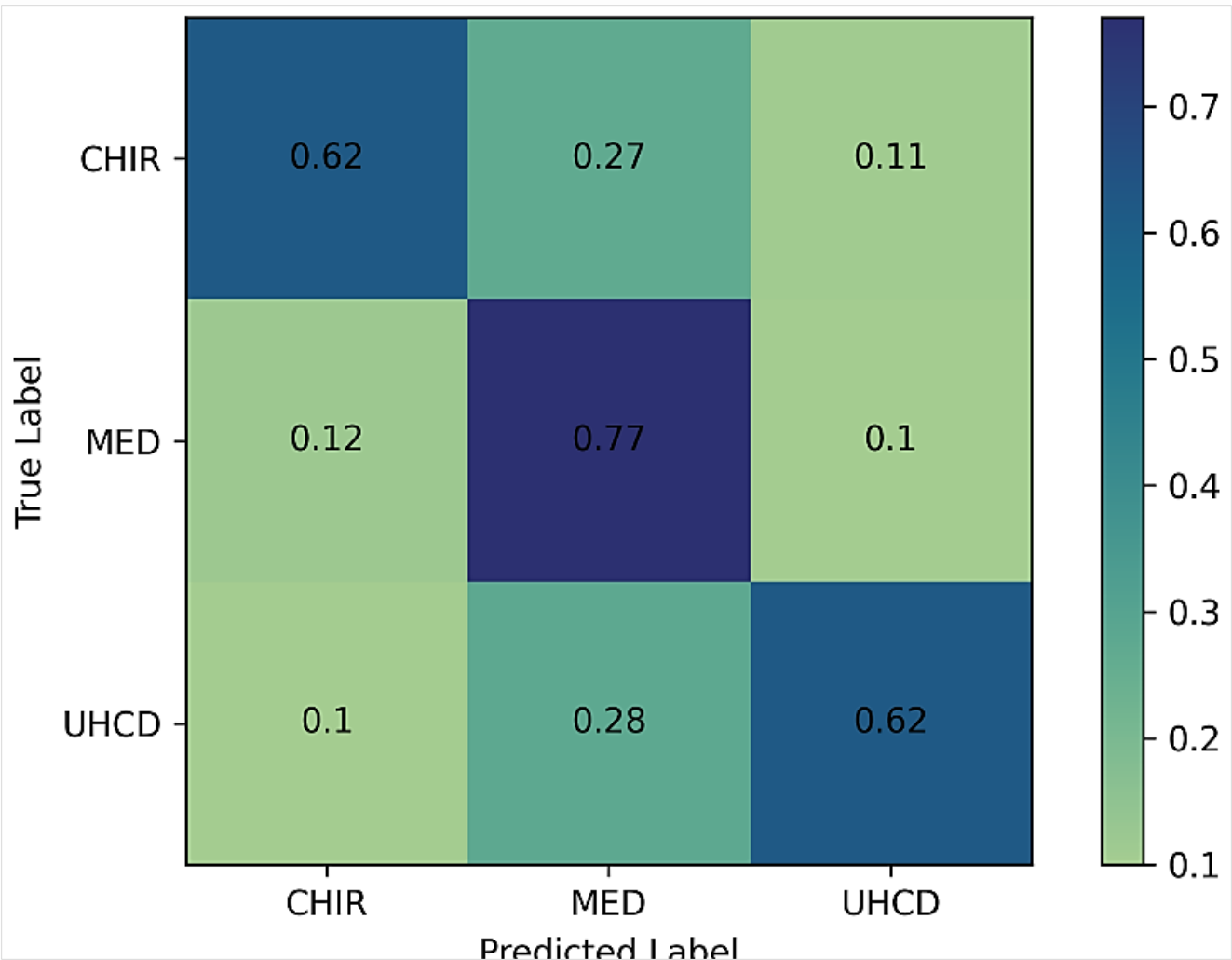
Text Preprocessing Pipeline



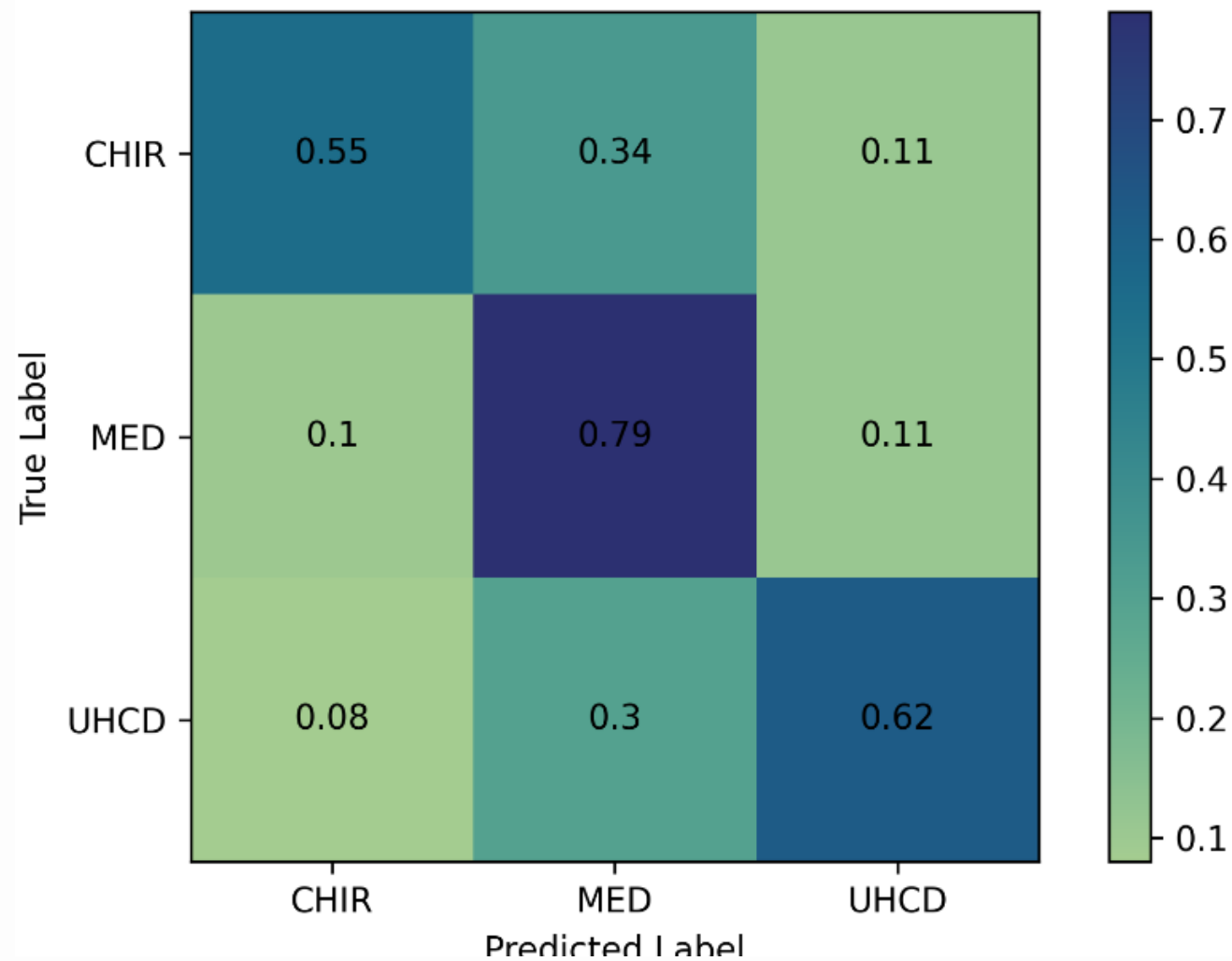
Model Architecture



Results (3-Fold CV)



Results (Model performance without text notes)



Conclusions

- ML-based approaches could be utilized to help provide an early prediction of medical specialties.
- The significance of our approach is the effective integration of standard data with free-text triage notes.
- Our empirical results robustly confirmed the positive impact of using such textual notes on the model performance.

Thank You!
