TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes

24th International Conference on Applications of Natural Language to Information Systems
NLDB 2019, University of Salford, United Kingdom

Tushaar Gangavarapu*, Aditya Jayasimha*, Gokul S Krishnan, and Sowmya Kamath S

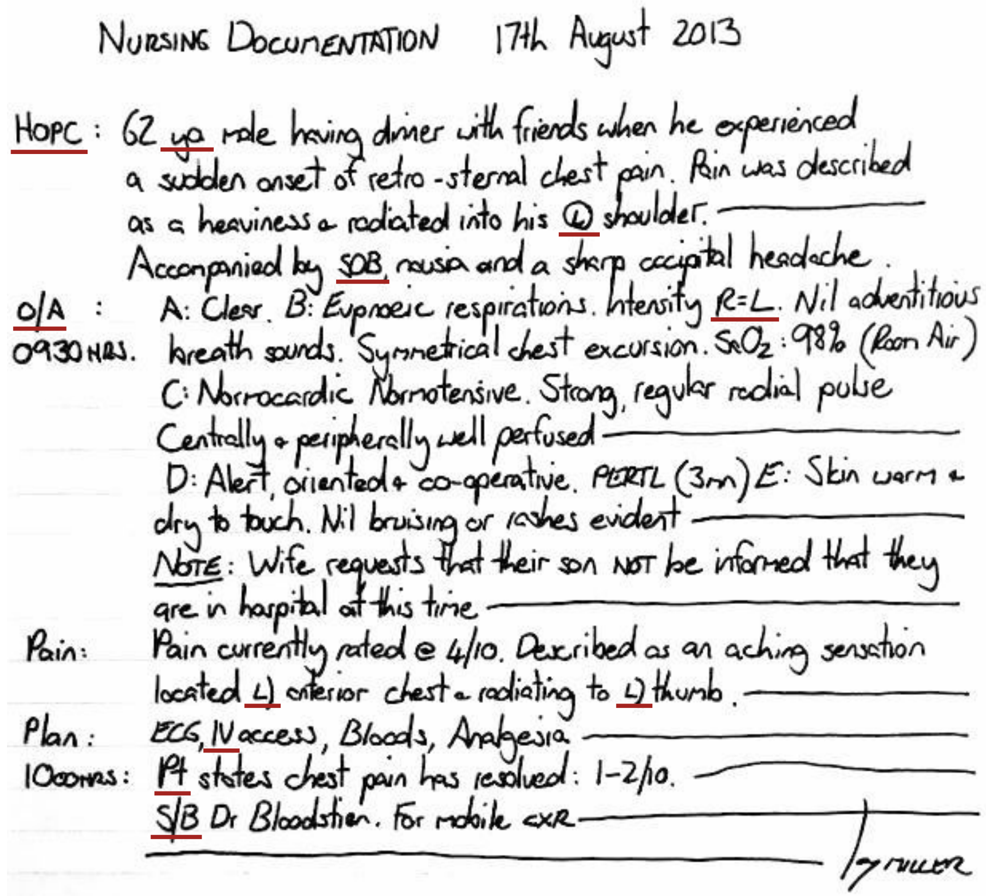
sowmyakamath@nitk.edu.in

Healthcare Analytics and Language Engineering (HALE) Lab
Department of Information Technology
National Institute of Technology Karnataka, Mangaluru, India

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Introduction

- Intensive care units are equipped with high medical resources[Huddar'16]
  - Risk assessment: etiology?
  - Staff and resources: bottleneck
  - Prevention: Postonset!
- Low rate of EMR adoption in developing countries: information loss?
- Clinical nursing notes (caregivers)
  - Subjective and objective assessments
  - Patient-centric and evidence-based
  - Raw, sparse, jargon, inconsistent, …



NURSING DOCUMENTATION 17th August 2013

HOPC: 62 yo male having dinner with friends when he experienced a sudden onset of retro-sternal chest pain. Pain was described as a heaviness & radiated into his (L) shoulder.
Accompanied by SOB, nausea and a sharp occipital headache.

O/A 0930 HRS: A: Clear. B: Eupnoeic respirations. Intensity R=L. Nil adventitious breath sounds. Symmetrical chest excursion. SaO_2: 98% (Room Air)
C: Normocardic. Normotensive. Strong, regular radial pulse
Centrally & peripherally well perfused
D: Alert, oriented & co-operative. PERTL (3mm) E: Skin warm & dry to touch. Nil bruising or rashes evident
NOTE: Wife requests that their son NOT be informed that they are in hospital at this time

Pain: Pain currently rated @ 4/10. Described as an aching sensation located (L) anterior chest & radiating to (L) thumb.

Plan: ECG, IV access, Bloods, Analgesia

1000HRS: Pt states chest pain has resolved: 1-2/10.
S/B Dr Bloodstien. For mobile CXR

A sample de-identified nursing note from critical care

[Huddar'16] V. Huddar *et al*. "Predicting complications in critical care using heterogeneous clinical data." IEEE Access 4 (2016): 7988-8001.

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Existing Methods

Work(s)	Methodology	Highlights	Remarks
Pirracchio et al. 2016 [1]	Super learner algorithm for the clinical task of ICU mortality prediction	Outperformed several severity scores	Did not benchmark against recent methods
Johnson et al. 2017 [2]	Case study on ICU mortality prediction highlighting the challenges in replicating results	Emphasis on the need to improve the way of performance reporting	Used feature sets rather than unstructured text
Harutyunyan et al. 2017 [3]	Multitask recurrent neural networks on four clinical prediction tasks	Promising results of deep learning approaches	Only benchmarked against LR and LSTM
Purushotham et al. 2018 [4]	Benchmarking of a suite of five clinical prediction tasks including ICD-9 code group prediction	Exhaustive benchmarking of deep architectures	Utilized only feature sets in the form of numerical assessments
Krishnan and Kamath 2018 [5]	ICU mortality prediction task using Word2Vec embeddings of ECG reports	Unsupervised data cleaning approach with clustering	Did not utilize deep neural architectures

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Existing Methods and Outcomes

- Systemized collection of patient data in the form of structured electronic medical records
  - Numerical assessments: lab investigations, medications, demographics, ...
  - Evidence-based precision medicine modeling?
  - Patient-specific?: unstructured clinical nursing notes
- Digitization: ~~manual or automated conversion of nursing notes to EMRs~~ Information loss
- AI-assisted modeling strategies: deep neural architectures in benchmarking of most state-of-the-art models?
- Voluminosity of nursing notes
  - Record of every observation made (subjective and objective)
  - Efficacy of underlying CDSS: information extraction and consolidation

Modeling similar notes!

- Multiple code assessment: manifold nature of disease symptoms and infections[Baumel'18]!

[Baumel'18] T. Baumel *et al*. "Multi-label classification of patient notes: case study on ICD code assignment." Workshops at 32nd AAAI Conference. 2018.

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Key Contributions

Aim: Development of a clinical decision support system to facilitate accurate risk assessment as ICD-9 code group prediction using unstructured clinical nursing notes

- ❖ Design of a fuzzy token-based similarity matching approach to effectively aggregate the voluminous unstructured clinical nursing notes of a patient, improving the efficacy of clinical decision making
- ❖ Utilization of vector space and coherence-based topic modeling approaches to extract the rich information available in unstructured clinical nursing notes and obtain optimal data representations
- ❖ Eliminating the dependency on the EMRs which is crucial in developing countries, through an effective approach that utilizes the abundantly available unstructured clinical text for disease prediction

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

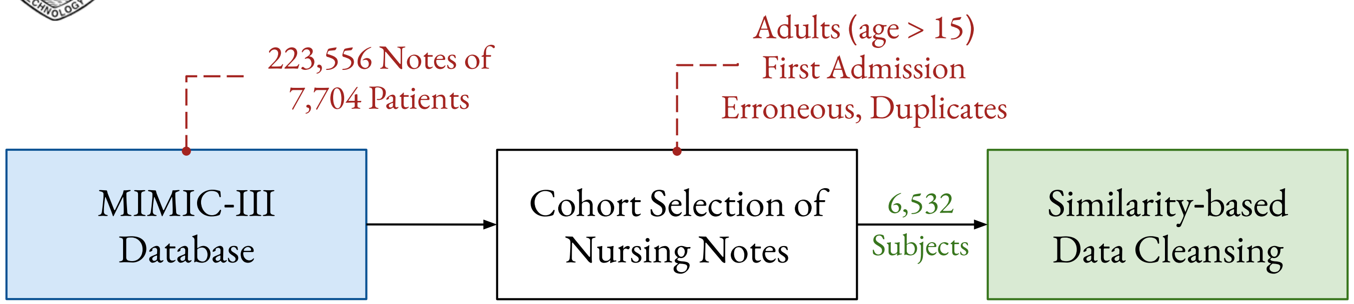
Proposed Methodology

MIMIC-III Database → Cohort Selection of Nursing Notes

- Diverse and comprehensive: de-identified healthcare database of over 40,000 ICU patients
  - 2,083,180 note events: 223,556 nursing notes of 7,704 patients
  - Four tables: *diagnosis_icd* (ICD-9 codes), *noteevents* (notes), *patients*, and *admissions* (age)
- Records of adults (age > 15) and first hospital admission were only retained[2][4][5]
  - Faster risk assessment: prediction with the earliest detected symptoms
  - Data cleaning: erroneous entries (*iserror* of *noteevents*) and duplicates → 6,532 patients

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Proposed Methodology



- Aggregation of nursing notes: voluminosity of nursing notes with near-duplicate entries
  - Monge-Elkan (ME) score: alternate names, clinical abbreviations, and medical jargon
  - Jaro score: spelling errors and normalization

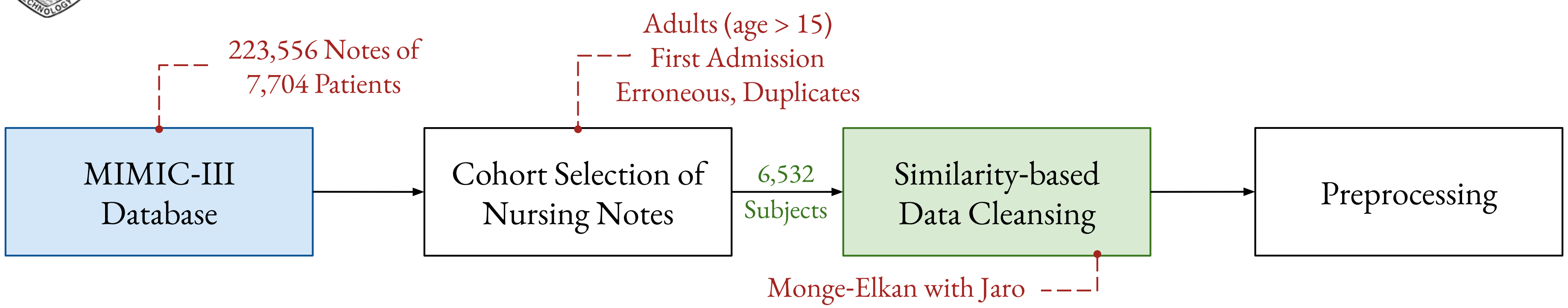
$$\mathrm{ME}_{\mathrm{Jaro}}(\eta_p, \eta_q) = \frac{1}{|\eta_p|} \sum_{i=1}^{|\eta_p|} \max \left\{ \mathrm{Jaro}(\mathcal{C}_i^{(p)}, \mathcal{C}_j^{(q)}) \right\}_{j=1}^{|\eta_q|} \quad \Bigg| \quad \mathrm{Jaro}(\mathcal{C}_m, \mathcal{C}_n) = \begin{cases} 0, & \text{if } c = 0 \\ \frac{1}{3}\left(\frac{c}{|\mathcal{C}_m|} + \frac{c}{|\mathcal{C}_n|} + \frac{2c-t}{2c}\right), & \text{otherwise} \end{cases}$$

  - Fuzzy-similarity threshold (0.825): merge or purge mechanism*

* Note that the merge or purge mechanism holds only for the nursing notes, and the ICD-9 codes across the corresponding nursing notes are always merged.

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

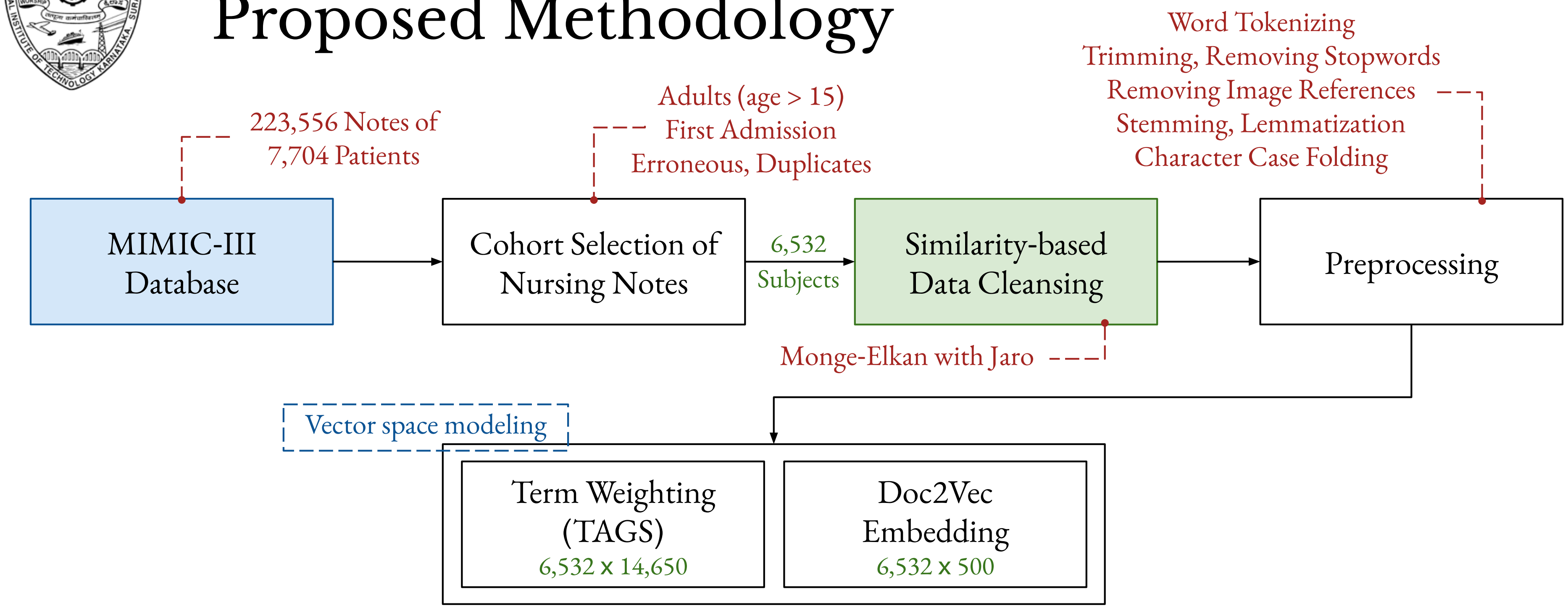
Proposed Methodology



- Data (text) normalization: standard canonical form
  - Space trimming, and removing punctuation and special characters
  - Tokenization: clinical nursing note to smaller words
  - Removing image references and character case folding
  - Stemming: suffix stripping, and lemmatization: conversion to base forms
- Overfit threshold: words in less than 10 nursing notes (computational complexity)

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

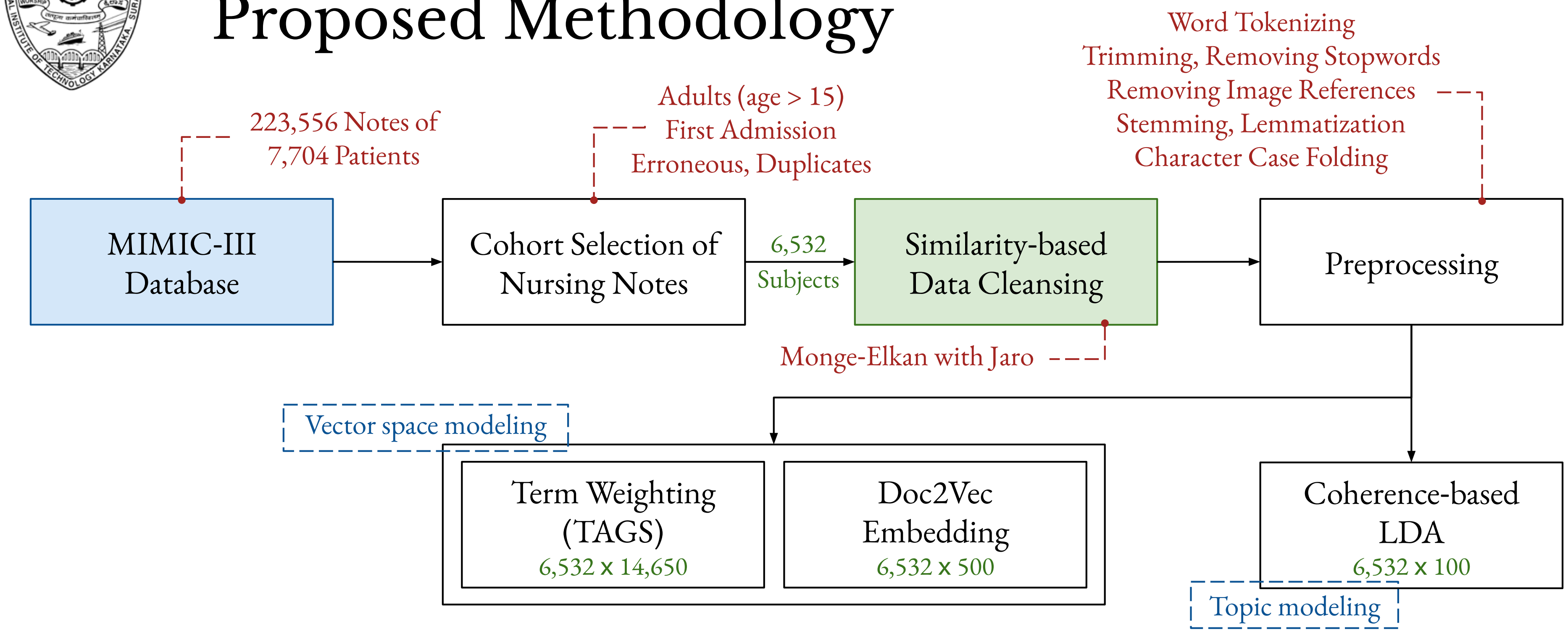
Proposed Methodology



- Vector space modeling: representation of each point in a multidimensional vector space
- Two variants: TAGS (term weighting scheme) and Doc2Vec (neural paragraph network)

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, “TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes,” International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

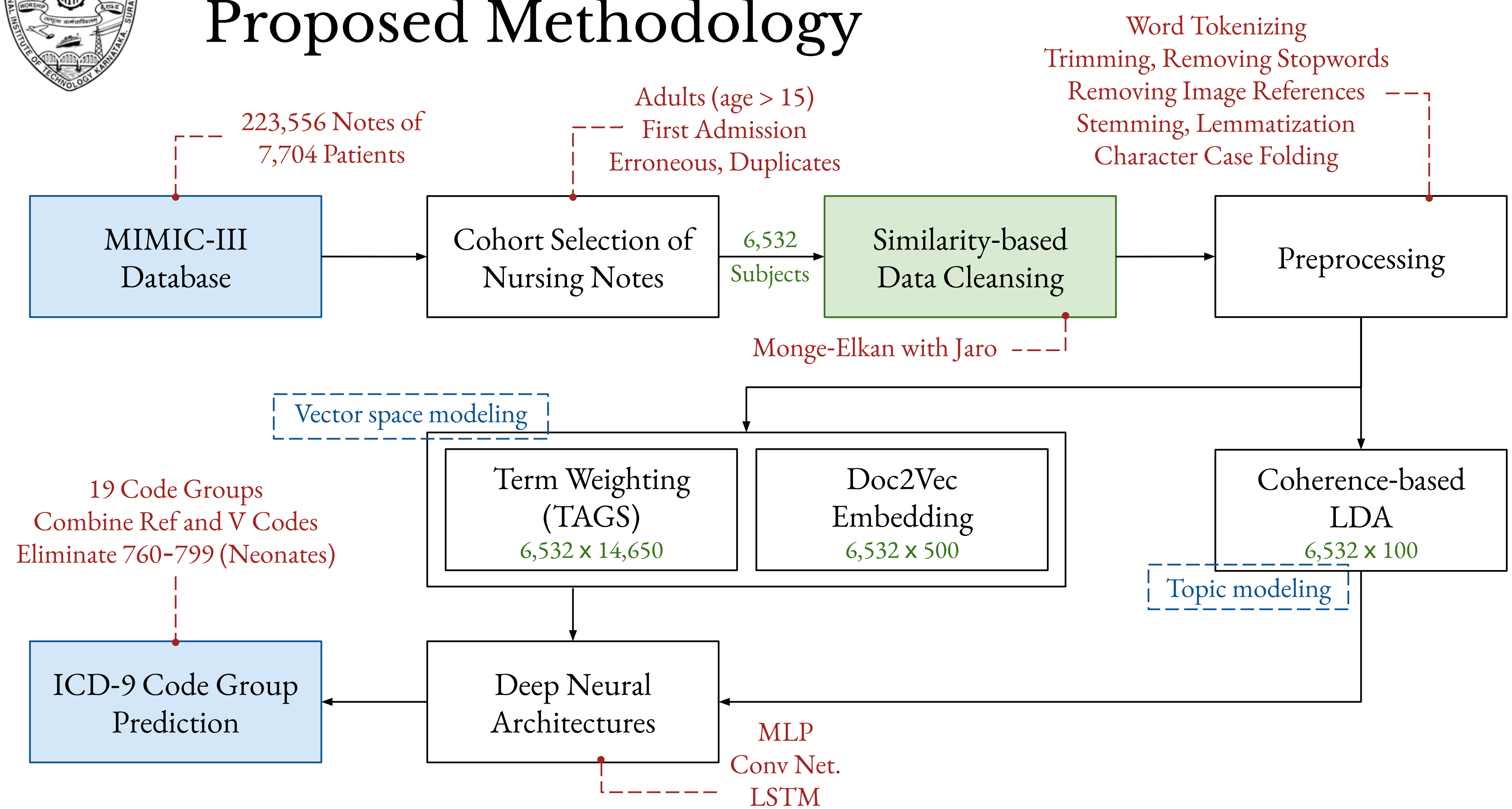
Proposed Methodology



- Topic modeling: finding a set of topics (collection of tokens) from a set of nursing notes
- Coherence scoring ensures higher guarantee on human interpretability

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Proposed Methodology



Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, “TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes,” International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Modeling of Clinical Concepts: Vector Space

- Vector space modeling: each nursing note is a point in multidimensional space ($d \ll |V|$)
- Term weighting scheme (transformation of bag-of-words): captures the syntax
  - Occurrence frequency: captures the importance of a term
  - Specificity: captures the rarity of a term

$$W_i^{(p)} = \begin{cases} \left(1 + log_2 f_i^{(p)}\right)\left(log_2 \frac{N}{|w^{(p)}|}\right), & \text{if } f_i^{(p)} > 0 \\ 0, & \text{otherwise} \end{cases} \rightarrow \{W_i^{(p)}\}_{i=1}^{|\mathbb{V}|}$$

Machine processable form

  - Curse of high dimensionality and sparsity!
- Doc2Vec (paragraph vectors): captures the semantics in the nursing notes (e.g., *bone* and *melanoma*)
  - Neural network: one shallow layer to learn distributed representations
  - Content-related measurement: syntax and semantics
  - Paragraph Vector-Distributed Memory: preserves word order[Le'14]

[Le'14] Q. Le and T. Mikolov. "Distributed representations of sentences and documents." International conference on machine learning. 2014.

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Modeling of Clinical Concepts: Topics

- Topic modeling: each nursing note is represented as a set of clinical topic clusters
- Latent Dirichlet Allocation: cluster analysis using Dirichlet distribution
  - Bayesian framework: documents, topics, and terms built on bag-of-words
  - Flat and soft probabilistic clustering: terms to topics, nursing notes to topics
  - Context-related assessment: captures syntax and latent semantics
- Topic Coherence: determine the optimal number of topics
  - Evaluation: human interpretability of the generated topics
  - Semantic similarity between high scoring terms
  - Confirmation metric: Normalized Pointwise Mutual Information (NPMI) score
  - NPMI: associations and collocations between terms[Aletras'13]
- Number of topics (100): vary number of topics in LDA model and compute coherence!

[Aletras'13] N. Aletras and M. Stevenson. "Evaluating topic coherence using distributional semantics." Proceedings of IWCS 2013–Long Papers. 2013.

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

ICD-9 Code Group Prediction

- International classification of diseases:
  - Diagnostic code group taxonomy
  - Epidemiology, research, billing, …
  - High granularity: categorization!
  - Disease-specific assessment
- Multi-label classification: groups
  - Manifold nature of symptoms
  - Aggregation: multiple code groups
- Diagnostic ICD-9 code groups (19):
  - 760 – 779: no records
  - Supplemental V and Ref. codes

http://tdrdata.com/ipd/ipd_SearchForICD9CodesAndDescriptions.aspx/

19 Code Groups

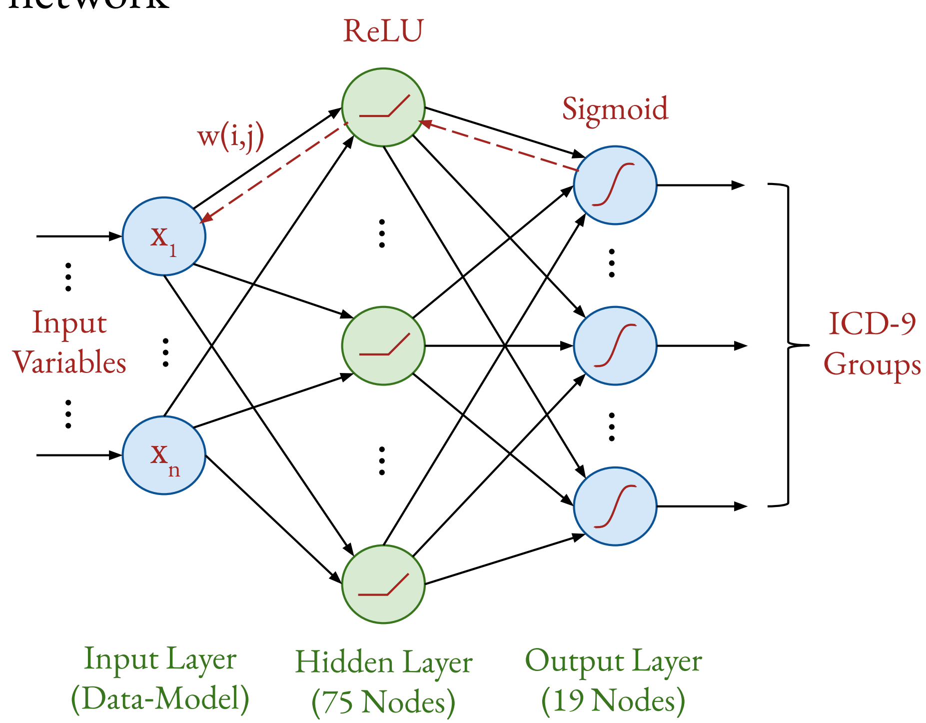
001 – 139	Infectious and Parasitic Diseases
140 – 239	Neoplasms
240 – 279	Endocrine, Nutritional, Metabolic, Immunity
280 – 289	Blood and Blood-Forming Organs
290 – 319	Mental Disorders
320 – 389	Nervous System and Sense Organs
390 – 459	Circulatory System
460 – 519	Respiratory System
520 – 579	Digestive System
580 – 629	Genitourinary System
630 – 677	Pregnancy, Childbirth, and the Puerperium
680 – 709	Skin and Subcutaneous Tissue
710 – 739	Musculoskeletal System and Connective Tissue
740 – 759	Congenital Anomalies
~~**760 – 779**~~	~~Conditions Originating in the Perinatal Period~~
780 – 789	Symptoms
790 – 796	Nonspecific Abnormal Findings
797 – 799	Ill-defined and Unknown Causes of Morbidity, Mortality
800 – 999	Injury and Poisoning
Ref. and V Codes	Reference and Supplemental V-Codes

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, “TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes,” International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

T. Gangavarapu *et al.*

Deep Neural Architectures

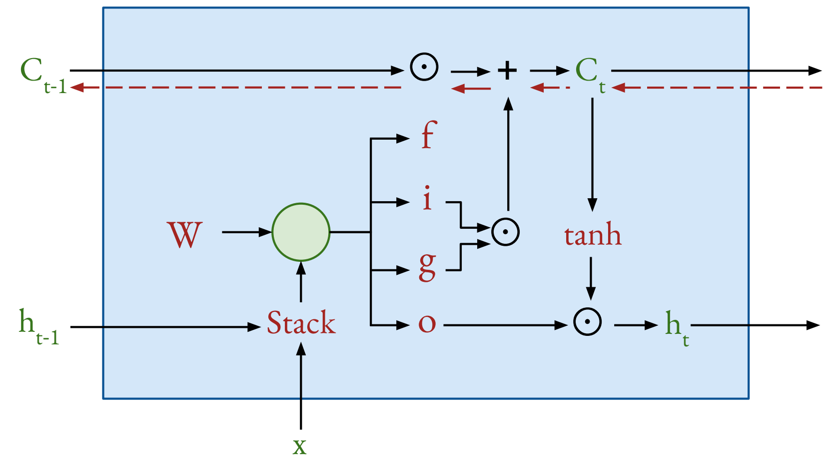
- Multi-layer perceptron: feed-forward neural network
  - Adaptive learning, fault tolerance, …
  - Layer-to-layer non-linear activation!
  - Backpropagation: optimal weights



Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, “TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes,” International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Deep Neural Architectures

- Multi-layer perceptron: feed-forward neural network
  - Adaptive learning, fault tolerance, ...
  - Layer-to-layer non-linear activation!
  - Backpropagation: optimal weights
- Long short term memory: long term dependencies
  - Vanishing gradient: additive interactions
  - Gated mechanism: *i*, *f*, *o*, and *g* gates
  - Multi-label classification: LSTM sigmoid

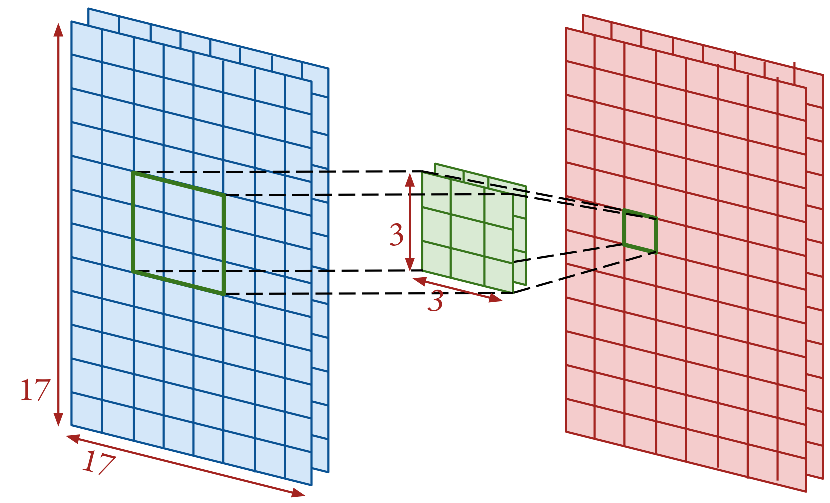


$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \quad \Bigg| \quad \begin{aligned} c_t &= f \odot c_{t-1} + i \odot g \\ h_t &= o \odot \tanh(c_t) \end{aligned}$$

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Deep Neural Architectures

- **Multi-layer perceptron**: feed-forward neural network
  - Adaptive learning, fault tolerance, …
  - Layer-to-layer non-linear activation!
  - Backpropagation: optimal weights
- **Long short term memory**: long term dependencies
  - Vanishing gradient: additive interactions
  - Gated mechanism: *i*, *f*, *o*, and *g* gates
  - Multi-label classification: LSTM sigmoid
- **Convolutional neural network**: minimal MLP
  - Hyper-parameter reduction
  - Classification: conv. flatten and dense



$$\mathbf{g}(f \cdot s^{(i)}_{k:k+w-1} + b) \leftarrow \{s^{(i)}_{1:w}, s^{(i)}_{2:w+1}, \ldots, s^{(i)}_{n-w+1:n}\}$$

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, “TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes,” International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Experimental Validation

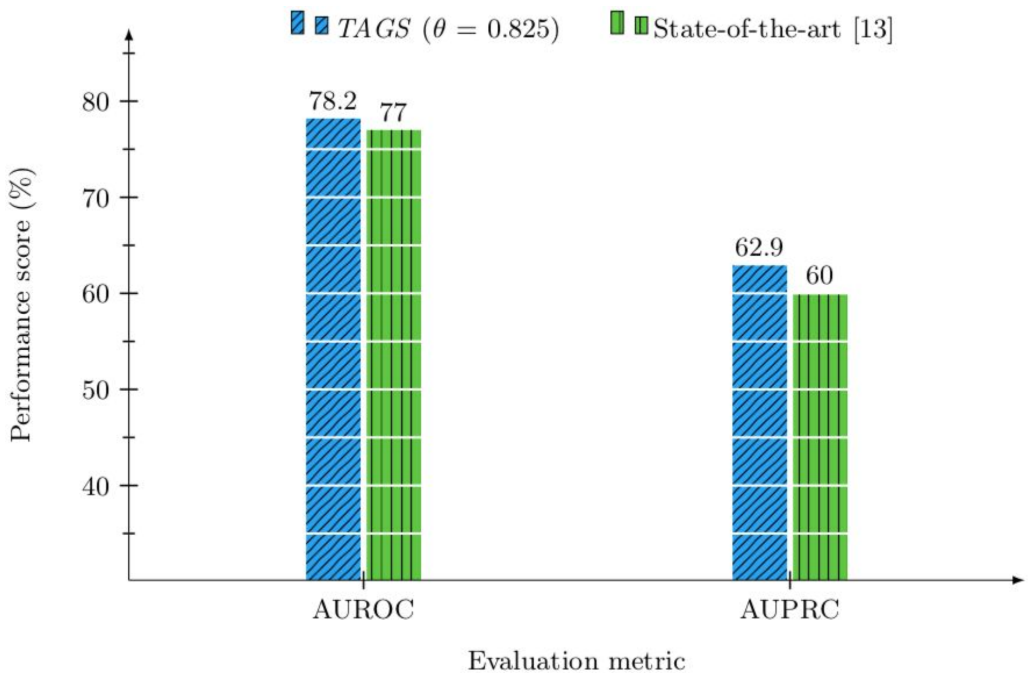
- Primary challenge: multi-label classification—pairwise comparison of actual and predicted!
- Evaluation (7): accuracy, MCC, F1, ranking loss*, coverage error*, AUPRC, and AUROC

Data Model	Classifier	Performance Scores				
		ACC	AUPRC	AUROC	MCC	F1
TAGS (6,532 x 14,650)	MLP	0.8130 ± 0.0005	0.6291 ± 0.0027	0.7738 ± 0.0013	0.5704 ± 0.0020	0.6803 ± 0.0024
	LSTM	0.7946 ± 0.0011	0.5990 ± 0.0014	0.7646 ± 0.0025	0.5365 ± 0.0027	0.6661 ± 0.0028
	CNN	0.8049 ± 0.0007	0.6153 ± 0.0031	0.7817 ± 0.0023	0.5594 ± 0.0022	0.6785 ± 0.0032
Doc2Vec (6,532 x 500)	MLP	0.7903 ± 0.0019	0.5914 ± 0.0016	0.7562 ± 0.0013	0.5212 ± 0.0032	0.6559 ± 0.0019
	LSTM	0.8005 ± 0.0017	0.6076 ± 0.0033	0.7600 ± 0.0010	0.5386 ± 0.0032	0.6655 ± 0.0022
	CNN	0.7737 ± 0.0012	0.5686 ± 0.0030	0.7433 ± 0.0019	0.4879 ± 0.0034	0.6381 ± 0.0028
LDA (6,532 x 100)	MLP	0.7905 ± 0.0017	0.5965 ± 0.0016	0.7497 ± 0.0017	0.5221 ± 0.0031	0.6397 ± 0.0027
	LSTM	0.7842 ± 0.0013	0.5865 ± 0.0012	0.7431 ± 0.0017	0.5078 ± 0.0014	0.6329 ± 0.0027
	CNN	0.8034 ± 0.0016	0.6181 ± 0.0011	0.7649 ± 0.0011	0.5542 ± 0.0022	0.6643 ± 0.0013

* Coverage error and (label) ranking loss are not presented here.

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Discussion



- Loss of information: EMR coding process from nursing notes
- TAGS: fuzzy aggregation by similarity
  - Patient centricity: nursing notes
  - Outperforms Doc2Vec and LDA
  - Class imbalance: AUPRC
  - MCC: balanced score
  - Seven benchmarking* metrics
- Outperforms state-of-the-art model[4]
  - AUPRC: 5%
  - AUROC: 1.55%

* This study employs five-fold cross-validation as the evaluation scheme.

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, “TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes,” International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Concluding Remarks

- Clinical nursing notes: treasure-trove of valuable patient-specific information
  - Patient-centric and evidence-based treatments
  - Accurate risk assessment as ICD-9 code group prediction
  - Modeling: rawness, sparsity, medical jargon, complex temporal nature, inconsistency, …
- Voluminosity of clinical nursing notes: similar or near-duplicate patient records
  - Aggregation: aids clinical decision making; reduces cognitive burden
  - Fuzzy similarity based nursing note matching: Monge-Elkan with Jaro score
- Existing models including state-of-the-art[4]: structured nature of EMRs
  - TAGS outperforms by 5% in AUPRC and 1.55% in AUROC
  - Performance assessment: seven benchmarking metrics (including AUPRC and AUROC)
- Research dependency on EMRs: TAGS—countries with low EMR adoption rates!

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

Acknowledgments

This work is funded by the Government of India's DST-SERB Early Career Research Grant (ECR/2017/001056) to Sowmya Kamath S

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

References

[1] Pirracchio, R. (2016). Mortality Prediction in the ICU Based on MIMIC-II Results from the Super ICU Learner Algorithm (SICULA) Project. In *Secondary Analysis of Electronic Health Records* (pp. 295-313). Springer, Cham.

[2] Johnson, A. E., Pollard, T. J., and Mark, R. G. (2017, November). Reproducibility in critical care: a mortality prediction case study. In *Machine Learning for Healthcare Conference* (pp. 361-376).

[3] Harutyunyan, H., Khachatrian, H., Kale, D. C., Steeg, G. V., and Galstyan, A. (2017). Multitask learning and benchmarking with clinical time series data. arXiv preprint arXiv:1703.07771.

[4] Purushotham, S., Meng, C., Che, Z., and Liu, Y. (2018). Benchmarking deep learning models on large healthcare datasets. *Journal of biomedical informatics*, 83, 112-134.

[5] Krishnan, G. S., and Kamath, S. S. (2018, June). A Supervised Learning Approach for ICU Mortality Prediction Based on Unstructured Electrocardiogram Text Reports. In *International Conference on Applications of Natural Language to Information Systems* (pp. 126-134). Springer, Cham.

[6] Bouma, G. (2009). Normalized (pointwise) mutual information in collocation extraction. *Proceedings of GSCL*, 31-40.

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, “TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes,” International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16

HALE Lab: AI-assisted NLP

Healthcare Analytics and Language Engineering (HALE) Lab

Research on Healthcare Analytics and Informatics, and also on Language and Linguistics using NLP

Headed by:

Dr. Sowmya Kamath S

Department of Information Technology

National Institute of Technology Karnataka, Mangaluru, India

Research Team:

Gokul S Krishnan

Karthik K Bhat

Veena Mayya

Tushaar Gangavarapu

Aditya Jayasimha

https://halelabnitk.github.io/

Please cite this article as: T. Gangavarapu, A. Jayasimha, G.S. Krishnan, and S.S. Kamath, "TAGS: Towards Automated Classification of Unstructured Clinical Nursing Notes," International Conference on Applications of Natural Language to Information Systems, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-23281-8_16