UNSUPERVISED PATIENT REPRESENTATIONS FROM CLINICAL NOTES WITH INTERPRETABLE CLASSIFICATION DECISIONS

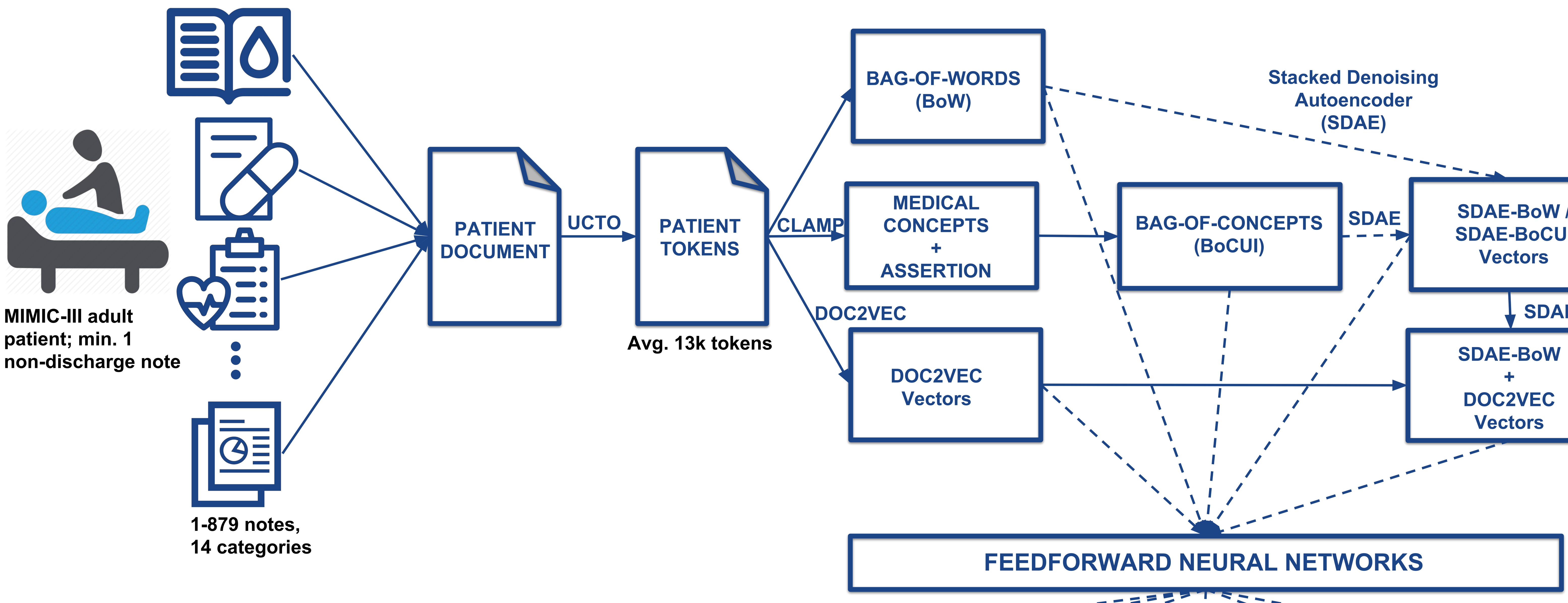
Madhumita Sushil[a,b,*], Simon Šuster[b], Kim Luyckx[a], Walter Daelemans[b]

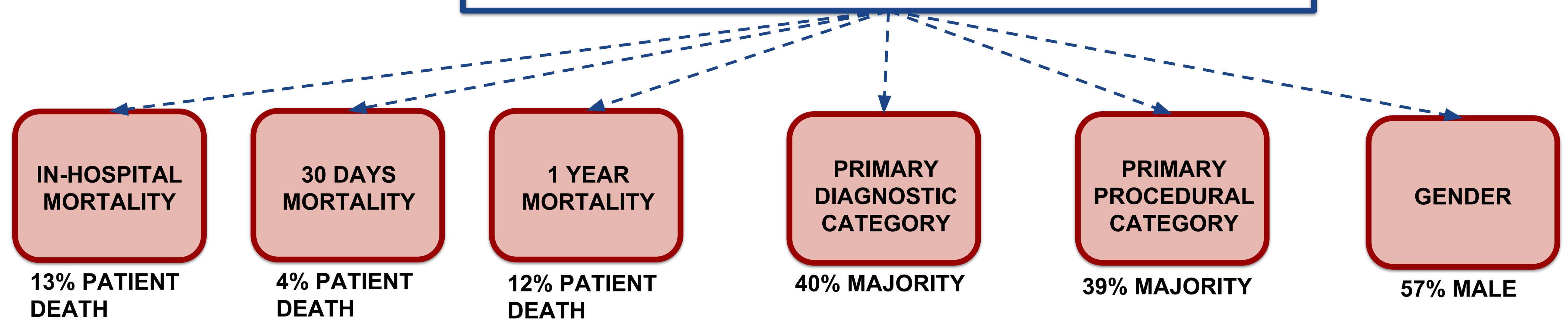
[a] Antwerp University Hospital (UZA), Edegem, Belgium

[b] Computational Linguistics and Psycholinguistics Research Center (CLiPS), University of Antwerp, Belgium

[*] madhumita.sushil@outlook.com

LEARNING TASK-INDEPENDENT PATIENT REPRESENTATIONS



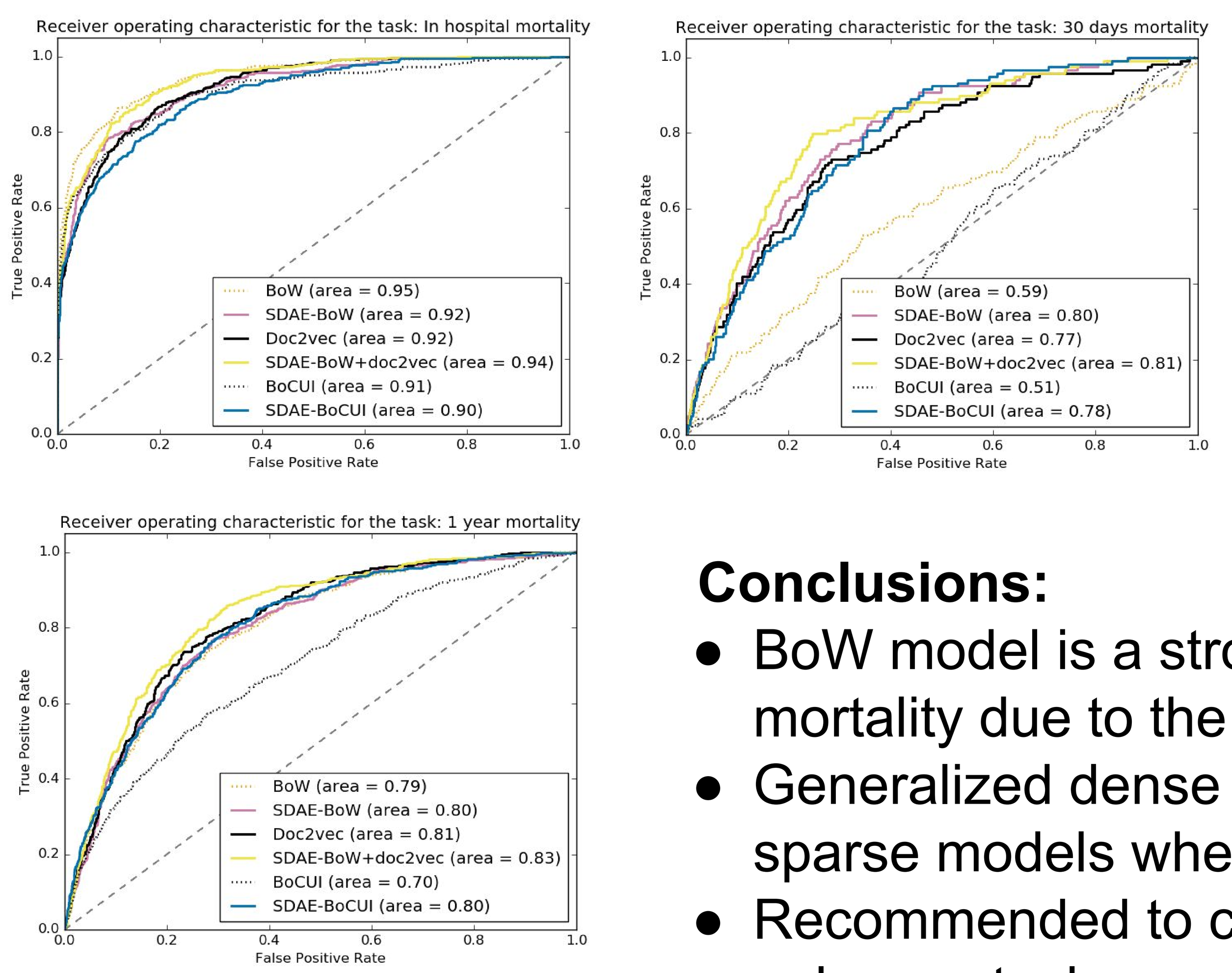


Goal: Similar patients ⇨ similar vectors, to facilitate patient cohort identification.

Novel Contributions:

- Evaluating if patient representation models are successful with only clinical notes as input.
- Analyzing if such models are transferable across tasks.
- Understanding the best and the worst encoded features in the SDAE representations.
- Extracting the most influential features for classification decisions.

REPRESENTATION EVALUATION



Approach	Pri_diag_cat (F-score-wt)	Pri_proc_cat (F-score-wt)	Gender (F-score-wt)
BoW	70.16	73.66	98.47
SDAE-BoW	65.00	67.46	87.75
Doc2vec	68.07	65.83	97.70
SDAE-BoW + Doc2vec	67.88	70.30	97.47
BoCUI	71.04	72.65	75.04
SDAE-BoCUI	66.47	67.77	62.45

Conclusions:

- BoW model is a strong baseline for all tasks except distant patient mortality due to the presence of strong lexical features.
- Generalized dense representation models significantly outperform sparse models when no. of positive instances is low (30 days mortality).
- Recommended to combine SDAE and doc2vec representations for unknown tasks.

FEATURE EXTRACTION

After **PRETRAINING** SDAE representations:

Compute **squared feature reconstruction error** after training the first SDAE layer, averaged across instances.

Finding: **High error correlation with frequency**.

CLASSIFICATION phase (input: SDAE representations):

Gradient of classification output w.r.t. original input, calculated using chain rule across networks.

Enables feature extraction for an arbitrary set of instances and output classes.

Finding: **Sensible, distinct features extracted for most tasks.**

In_hosp	30_days	1_year	Pri_diag_cat	Pri_proc_cat	Gender
vasopressin	leaflet	magnevist	numeric_val	numeric_val	woman
pressors	structurally	signal	previous	no	female
focused	pacemaker	decisions	rhythm	of	she
dnr	sda	periventricular	no	enzymes	man
dopamine	periventricular	embolus	flexure	extubated	he
acidosis	excursion	underestimated	dementia	rhythm	male
levophed	non-coronary	calcified	brbpr	and	her
pressor	dosages	screws	of	the	his
cvvhd	microvascular	rib	sinus	vent	wife
cvvh	left-sided	shadowing	for	uncal	uterus
emergency	chronic	gadolinium	to	mso	him

Acknowledgements: This research was carried out within the **Accumulate** project (www.accumulate.be), funded by the government agency Flanders Innovation and Entrepreneurship (VLAIO), Belgium [grant no. 150056].