

2018 n2c2 shared task

Track 1 Wrap-up

Amber Stubbs, Michele Filannino, Ozlem Uzuner

Assistant Professor, Simmons University

November 2, 2018

Aggregate metrics - micro f1, all runs

	Track 1
Minimum	0.2117
Maximum	0.91
Average	0.799
Standard deviation	0.116
Median	0.8227

Top 10 teams

* = Presentation

= Poster

Team	System Description
MedUniGraz	Rules + regular expressions
Univ of Michigan	Hybrid; external resources (MetaMap, cTAKES, and RxNORM) + rules
Sorbonne Université	Hybrid; rule-based + terminology-based + semi-supervised ML
Med Data Quest	Rule-based ; “bottom-up”; separate modules for different aspects of the files
Cincinnati CHMC	Hybrid; rules + external resources + ML
ASU	Hybrid; CLAMP + rules + some SVM
UNSW / NCI	Rule-based ; consulting with MDs
Harbin IT	Ensemble of CNN and CNN-highway-LSTM
University of Utah	Rule-based , with a trie-based hash rule
NTTMUNSW	rules + SVM with polynomial kernel

Overall stats for top 10 systems

Rule-based: 4

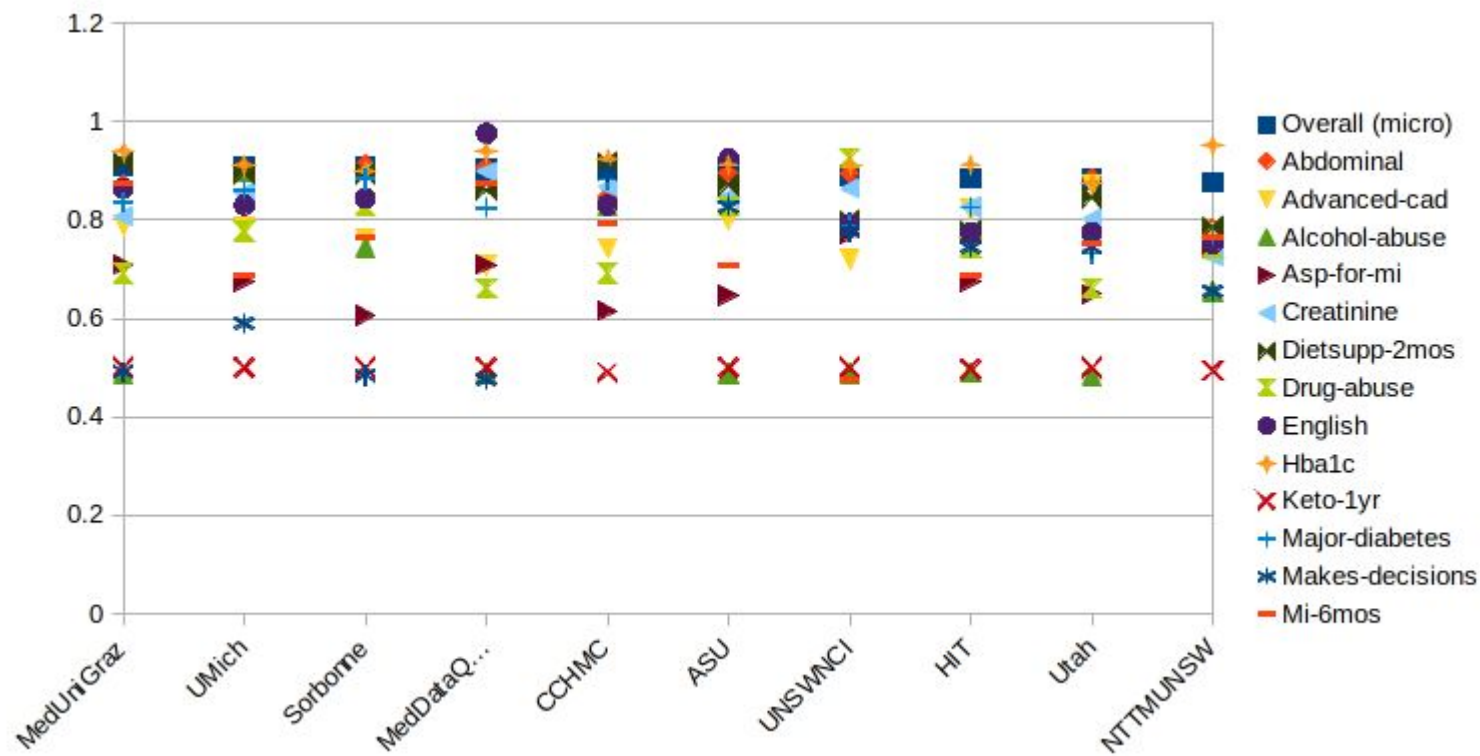
Hybrid: 6

Machine learning: 0

Used MDs: 5/10

External resources (aggregate): UMLS, cTAKES, MetaMap, NIH resources, Wikipedia, Web resources, MIMIC III, Dietary Supplement Label Database, Ontologies, medlineplus dietary suppl, med lists for disease, webmd, medlineplus

Team results by criteria (Micro F1)



Overall

High-scored categories

- Lab tests: HbA1c value between 6.5 and 9.5%, Serum creatinine > upper limit of normal
- Majority of cases are “met”: Patient must speak English

Low-scored categories:

- Few examples in training: Diagnosis of ketoacidosis in the past year
- Difficult to categorize: Patient must make their own medical decisions, Current alcohol use over weekly recommended limits
- Requires reasoning: Use of aspirin to prevent myocardial infarction

Gold standard statistics

These are the original numbers, not the numbers from the corrections found during the shared task

Criterion	Met	Not met
Abdominal	107	181
Advanced CAD	170	118
Alcohol abuse	10	278
Aspirin for MI	230	58
Creatinine	106	182
Diet supplement w/in 2 months	149	139
Drug abuse	15	273
Speaks English	265	23
HBA1C	102	186
Ketoacidosis within 1 year	1	287
Major diabetes	156	132
Makes own decisions	277	11
MI in 6 months	26	262

JAMIA special issue

https://academic.oup.com/jamia/pages/cfp_cohort_selection_for_clinical_trials

Special Focus Issue on Cohort Selection for Clinical Trials

Papers about Track 1 systems, and other relevant research

January 16, 2019 - Paper submission deadline

“All manuscripts, including those from the participants in the 2018 n2c2 Shared-Task and Workshop on Cohort Selection for Clinical Trials, should clearly present the novelty of the work and contribution to the field.”

Data availability

Data will be available for research outside of the shared task in November 2019, from <https://n2c2.dbmi.hms.harvard.edu/track1.php>

Thank you!

Organizing committee:

- Ozlem Uzuner, co-chair, George Mason University
- Amber Stubbs, co-chair, Simmons University
- Michele Filannino, co-chair, MIT
- Kevin Buchan, SUNY at Albany
- Susanne Churchill, Harvard Medical School
- Isaac Kohane, Harvard Medical School
- Hua Xu, UTHealth
- Ergin Soysal, UTHealth

National Library of Medicine, National Institutes of Health, NIH NLM 5R13LM011411, PI: Ozlem Uzuner



HARVARD
MEDICAL SCHOOL

DEPARTMENT OF
Biomedical Informatics



Volgenau School
of Engineering