Explainable Prediction of Medical Codes from Clinical Text

An, Tran Cong Viet   Quan, Nguyen The

Department of Information Technology, University of Engineering and Technology, Vietnam National University

April 6, 2022

Table of Content

Preamble

This slide is adapted from James Mullenbach et al. (2018). "Explainable Prediction of Medical Codes from Clinical Text". In: *CoRR* abs/1802.05695. arXiv: 1802.05695. URL: http://arxiv.org/abs/1802.05695.

You can get this slide at **https://antran22.github.io/files/medical-code.pdf**

Clinical Notes

- Notes produced by clinician.
- Written or dictated text outlining the interaction with patients.
- Detailed, Accurate.

A Sample Clinical Note

Admission Date :
⟨ deidentified ⟩
Discharge Date :
⟨ deidentified ⟩
Date of Birth :
⟨ deidentified ⟩ Sex :
F
Service :
SURGERY
Allergies :
Patient recorded as having No Known Allergies to Drugs
Attending :
⟨ deidentified ⟩
Chief Complaint :
Dyspnea
Major Surgical or Invasive Procedure :
Mitral Valve Repair
History of Present Illness :
Ms. ⟨ deidentified ⟩ is a 53 year old female who presents after a large bleed rhythmically lag to 2 dose but the patient was brought to the Emergency Department where he underwent craniotomy with stenting of right foot under the LUL COPD and transferred to the OSH on ⟨ deidentified ⟩ .
The patient will need a pigtail catheter to keep the sitter daily .

Figure: A sample clinical code

Medical Codes

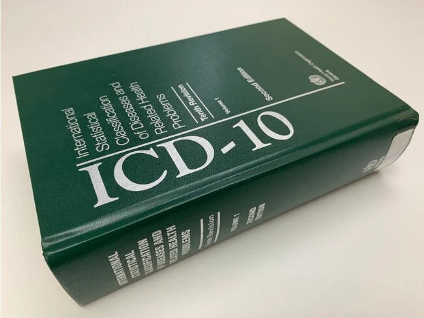


Figure: Dictionary of ICD-10, a Clinical Code Standard WHO

- **Standardised** Codes
- Each code maps to **disease** or **medical procedure**.

Medical Codes Standard: ICD

International Classification of Disease
Maintained by World Health Organization.

- **ICD-9**: Developed from 1975, adopted from 1978.
- **ICD-10**: Developed from 1983, adopted from 1994. Still widely used.
- **ICD-11**: Developed from 2007, released at the start of 2022.

This article focus on **ICD-9** standard, because the dataset is in this standard

Medical Codes Standard: CPT

Current Procedural Terminology
Assign a code for each task and service that a healthcare worker may perform.

Why do we need Medical Codes

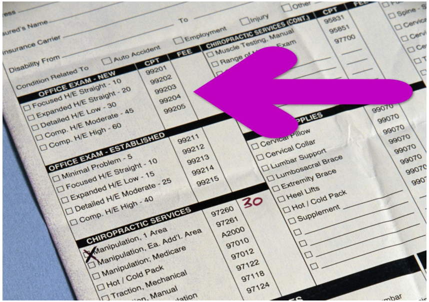


Figure: A Medical bill. The arrow is pointing at CPT codes for Medical Procedures

- Billing
- Pharmatical prescription
- Modeling patient state [1]

[1] Choi et al. 2016.

Electronic Health Record

Digital Record of Patients' Health.

- Medical History, Allergies.
- Current Diagnostic.
- Medications.

Can be quickly accessed by authorized personnels (doctors, healthcare workers).

EHR, Medical Codes, Clinical Notes

- EHR should store *structured data* (Medical Codes) for efficient lookup.
- Clincal Notes are *unstructured*.

$\Longrightarrow$ Need an operation for converting **Clinical Notes** to **Medical Codes**

The Note-to-Code Process

1. Clinical Notes.
2. Extract phrases that contains informations about medical problem.
3. Map phrases to Medical Codes (ICD-10).

Manual Coding

Figure: A manually indexed code book for medical coding

- Very laborious process.
- Error-prone [2].

[2] Birman-Deych et al. 2005.

Automatic Coding

In research since 1990s [3]

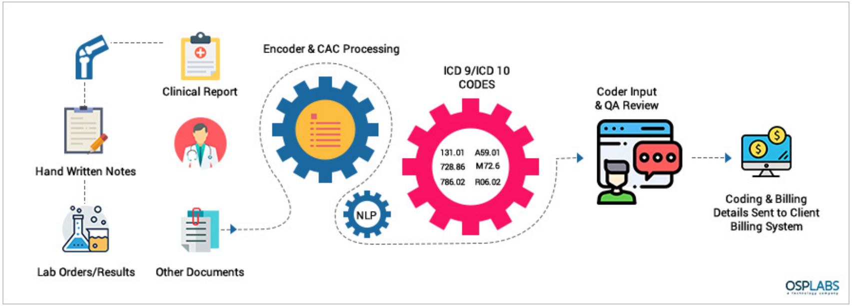


Figure: The process of automated clinical coding

[3] Lima et al. 1998.

The Clinical Coding problem from NLP perspective

- Multi-label classification problem
- Label space size: 14,000
- Input are long, **semi-structured** documents.

The MIMIC dataset

- Open-access, de-identified health-related data .
- 47k documents.
- Released by Johnson et al. 2016.

Properties of MIMIC documents

- **Loosely structured**.
- **Long**: Post-processed document length: Median 1,341.
- Use ICD-9 medical code as labels.
- Number of labels in each document: Median 14.

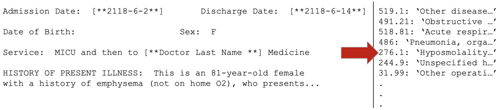
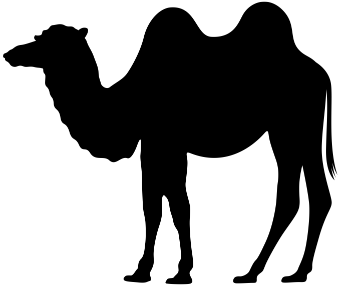


Figure: MIMIC dataset example

CAML

Convolutional **A**ttention for **M**ulti-**L**abel Classification
Focus on highly informative short span of text.

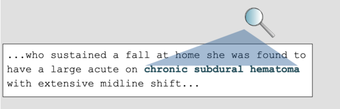


Desired Property of CAML: Precise Focus

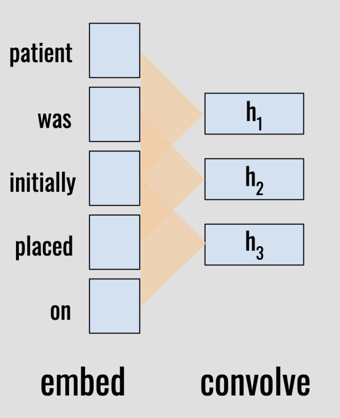
Mullenbach et. al conjectured that informations about diseases are condensed in short, informative text span.

- Different phrases for different diseases.

$\Longrightarrow$ Model should focus in short phrases.



Convolution Layer



- The Convolution Layer is useful for focusing on short phrases.
- Mullenbach et. al choose a kernel size of 10; i.e., focus on 10 words phrases in the text.

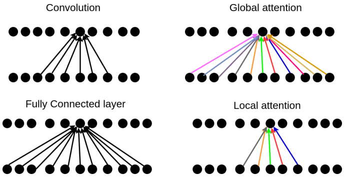
Not every word is equally important

Consider this phrase describing a symptom of *Foreign body in left main bronchus* - ICD-9 code: **934.1**.

> *...bronchoscopy performed showing large mucus plug on the left..*

Only **bronchoscopy**, **mucus**, **plug**, **left** are informative.

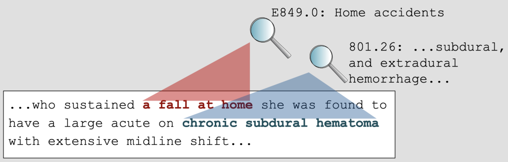
What is Attention Mechanism



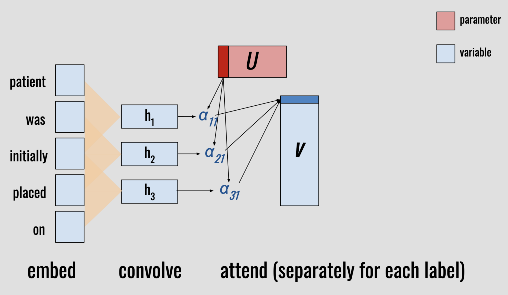
- A learned, auxiliary model for NLP models.
- Learn how to assign weights for input tokens (different focus on tokens)
- Proposed by: Bahdanau et al. 2015 & Luong et al. 2015

Desired Property of CAML: Treat labels individually

Each label should focus on different phrases.



Attention for Individual Label



Each column of Attention Model **U** is learnt separately for each label

Each row of Matrix **V** is calculated for each label.

Classification

A Simple Linear layer β with Softmax activation is used to calculate the probabilities for each labels.

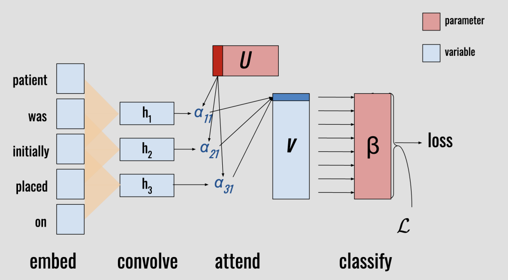


Figure: Model Structure of CAML

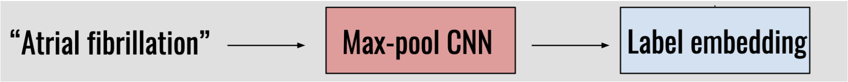
Some labels are very similar

The descriptions of two similar codes.

- **250.00**: "Diabetes mellitus without mention of complication, type II or unspecified type, *not stated* as uncontrolled"
- **250.02**: "Diabetes mellitus without mention of complication, type II or unspecified type, uncontrolled"

Get more information about labels

An embedding for labels will be trained separately on labels' descriptions.



Regularization of Classification Layer

L2 loss function to regularize β by the **Label Embedding**.
$\Longrightarrow$ Similar labels has similar parameters in β.

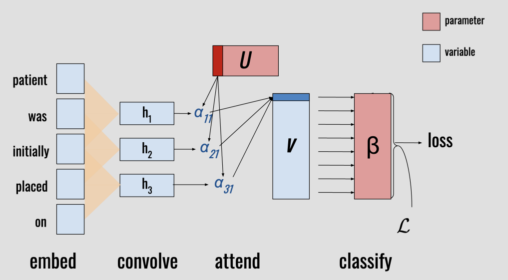


Figure: Model Structure of DR-CAML (Description Regularized CAML)

To be Done

Citation I

- Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio (2015). "Neural Machine Translation by Jointly Learning to Align and Translate". In: ed. by Yoshua Bengio and Yann LeCun. URL: http://arxiv.org/abs/1409.0473.
- Birman-Deych, Elena, Amy D Waterman, Yan Yan, David S Nilasena, Martha J Radford, and Brian F Gage (May 2005). "Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors". en. In: *Med Care* 43.5, pp. 480–485.
- Choi, Edward, Mohammad Taha Bahadori, Andy Schuetz, Walter F. Stewart, and Jimeng Sun (2016). "RETAIN: Interpretable Predictive Model in Healthcare using Reverse Time Attention Mechanism". In: *CoRR* abs/1608.05745. arXiv: 1608.05745. URL: http://arxiv.org/abs/1608.05745.
- Johnson, Alistair E W et al. (May 2016). "MIMIC-III, a freely accessible critical care database". In: *Scientific Data* 3.1, p. 160035.

Citation II

- Lima, Luciano R. S. de, Alberto H. F. Laender, and Berthier A. Ribeiro-Neto (1998). *A Hierarchical Approach to the Automatic Categorization of Medical Documents*. CIKM '98. Bethesda, Maryland, USA: Association for Computing Machinery, pp. 132–139. ISBN: 1581130619. DOI: 10.1145/288627.288649. URL: https://doi.org/10.1145/288627.288649.
- Luong, Minh-Thang, Hieu Pham, and Christopher D. Manning (2015). "Effective Approaches to Attention-based Neural Machine Translation". In: *CoRR* abs/1508.04025. arXiv: 1508.04025. URL: http://arxiv.org/abs/1508.04025.
- Mullenbach, James, Sarah Wiegreffe, Jon Duke, Jimeng Sun, and Jacob Eisenstein (2018). "Explainable Prediction of Medical Codes from Clinical Text". In: *CoRR* abs/1802.05695. arXiv: 1802.05695. URL: http://arxiv.org/abs/1802.05695.