Deep Learning Approaches for Predicting Glaucoma Progression Requiring Surgery Using Electronic Health Records Data and Natural Language Processing

Sophia Y Wang, MD[1]; Benjamin Tseng[1]
[1]Byers Eye Institute, Stanford University, Palo Alto, CA

Introduction

- Advances in artificial intelligence have enabled predictive models in glaucoma, including a previously published logistic regression model predicting glaucoma progression to surgery with **Area Under Curve (AUC) of 0.67**.[1] However, uncertainty exists regarding how to integrate the wealth of information in free-text clinical notes.
- Many **clinical details are in the free-text clinical progress** notes in the electronic health records (EHR), which are difficult to access and compute over.
- The use of **neural word embeddings** is a **natural language processing** technique where words are mapped into numeric vectors, such that word "meaning" is encoded within the vector space. Neural word embeddings provide an approach to integrating text into predictive models.

Purpose

- To build and evaluate **deep learning (DL) models** using ophthalmology domain-specific neural word embeddings to represent clinical notes, in order **to predict glaucoma progression requiring surgery**.
- To compare the performance of models that integrate free-text notes with those that used only structured input data.

Methods

- Cohort Building:

01 Stanford Research Repository with EHR from 2008-2020
- Patients with who underwent incisional glaucoma surgery (CPT)
- Patients with ≥ 2 instances of a glaucoma diagnosis who did not get surgery

02
- N=1298 surgical patients
- N=5050 nonsurgical glaucoma patients

03
- N=748 surgical patients who had at least 120 days of follow-up without getting surgery
- N=3764 nonsurgical patients with at least 120 days of follow-up

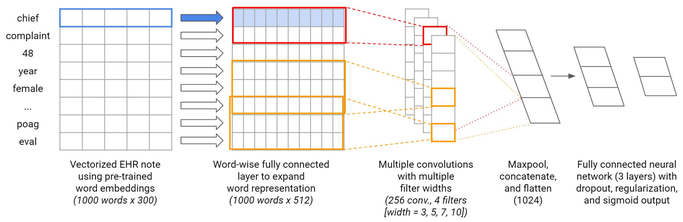
- Feature Engineering
  - Unstructured (Text):
    - Identified first 3 notes from within first 120 days of follow-up.
    - All notes lower-cased, tokenized (split into separate words), and stopwords (a, and, the, etc.) removed
    - Mapped to 300-dimensional neural word embeddings customized for ophthalmology, pre-trained on PubMed ophthalmology abstracts.
  - Structured (total 361 features):
    - Boolean: Billing codes (ICD and CPT), medications. All near zero variance features removed.
    - Numeric: eye exam information for both eyes[2], summarized with high, low, most recent, mean, missing value indicators

Deep Learning Model Architectures

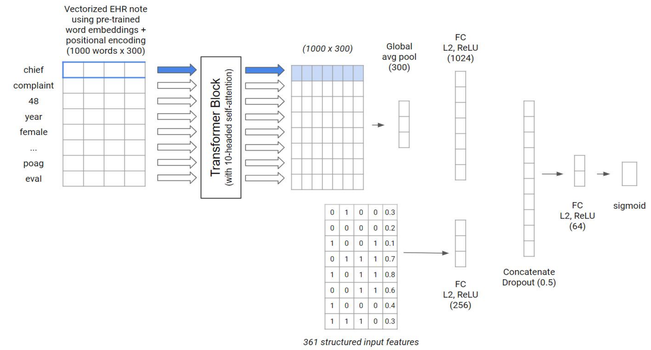
Structured Model

Structured features → 1024D layer (ReLU, L2 regularization) → Dropout (0.5) → 64D layer (ReLU, L2 regularization)→ sigmoid output

Text Model

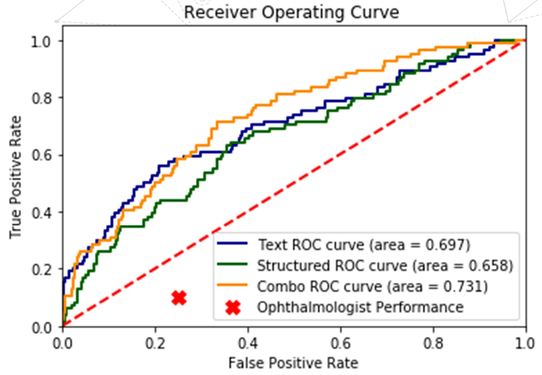


Combined Text and Structured Model

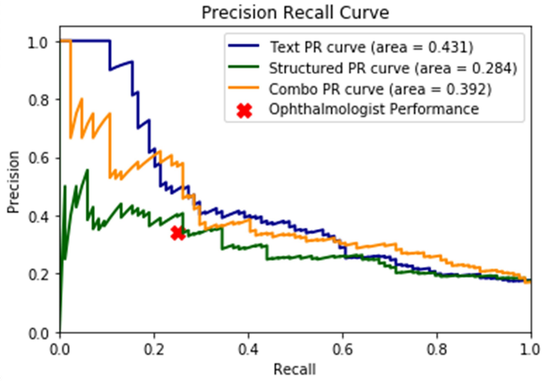


Classification Target For All Models:
Patients who progressed to incisional glaucoma surgery (16.6% of population)

Results



- Combined model slightly outperforms others in AUROC
- All models outperform ophthalmologist, who has low sensitivity but relatively high specificity



- Text model has slightly better AUPRC

	F1	Sensitivity (Recall)	Specificity	PPV (Precision)	NPV	Accuracy	Threshold
Ophthalmologist Predictions	0.29	0.25	**0.90**	**0.34**	0.85	**0.79**	-
Structured Model	0.34	0.69	0.53	0.23	0.89	0.56	0.15
Text-Only Model	**0.42**	0.56	0.77	0.33	0.90	0.74	0.20
Combined Model	0.40	**0.77**	0.57	0.27	**0.93**	0.60	0.15

Conclusions and Discussion

- Using word embeddings to represent clinical notes, **deep learning models were able to predict whether glaucoma patients would need glaucoma surgery in the future,** at a performance level better than an ophthalmologist review of the same notes
- Models incorporating text performed better than models using only structured (non-free-text) data.
- **Limitations**:
  - Observational, single center study
  - Imaging information not directly incorporated (only interpretations, if written in the clinical notes)
  - One prediction provided based on baseline data, rather than continuous predictions updated with each new piece of data
- **Clinical relevance**:
  - Predictive models can be helpful in clinical decision support, or in automatically identifying high-risk patients for clinical trials
  - However, performance still must be greatly improved before any deployment
- **Future work:**
  - Expand to multiple centers (if interested in collaboration, please contact me).
  - Integrate imaging into predictive models as another modality of data
  - Use of more sophisticated representation methods for text, such as transformer-based models
  - Use of named entity recognition systems to produce features from the clinical text
  - Investigations of performance in subgroups of patients (e.g. by physician, race/ethnicity)

References

1. Baxter, Sally L., Charles Marks, Tsung-Ting Kuo, Lucila Ohno-Machado, and Robert N. Weinreb. 2019. "Machine Learning-Based Predictive Modeling of Surgical Intervention in Glaucoma Using Systemic Data From Electronic Health Records." *American Journal of Ophthalmology* 208 (December): 30–40.
2. Wang, Sophia Y., Suzann Pershing, Elaine Tran, and Tina Hernandez-Boussard. 2020. "Automated Extraction of Ophthalmic Surgery Outcomes from the Electronic Health Record." *International Journal of Medical Informatics*. https://doi.org/10.1016/j.ijmedinf.2019.104007.

Supported by: NIH National Eye Institute P30-026877; Research to Prevent Blindness Career Development Award Departmental Grant

Email sywang@stanford.edu for more information.

Authors have no conflicts of interest to disclose.