Developing an AI Model to Identify Malignancy Results on Imaging Reports to Delay Release to the Electronic Patient Portal

Lucy Chow, Paul Lukac, Pin-Chieh Wang, Adam McCue, Vu Vu, Mojtaba Heidarysafa, Kamran Kowsari, Mark McArthur, Eric Cheng

2023 Radiological Society of North America Conference

November 26-30, 2023

Lucy Chow, MD
Assistant Clinical Professor, Radiology
Physician Informaticist
lchow@mednet.ucla.edu

UCLA Health
Information Technology

OHIA | Office of Health Informatics & Analytics

Summary of the Laws on Imaging Result Release

21st Century CURES Act	SB1419
Federal law	California state law
Took effect 10/2020	Enacted 10/2022, Took effect 01/2023
Immediate release of results to electronic patient portal	Imaging scan results that **reveal a new or recurrent malignancy** shall not be released to the patient portal until discussion with the patient by a healthcare professional

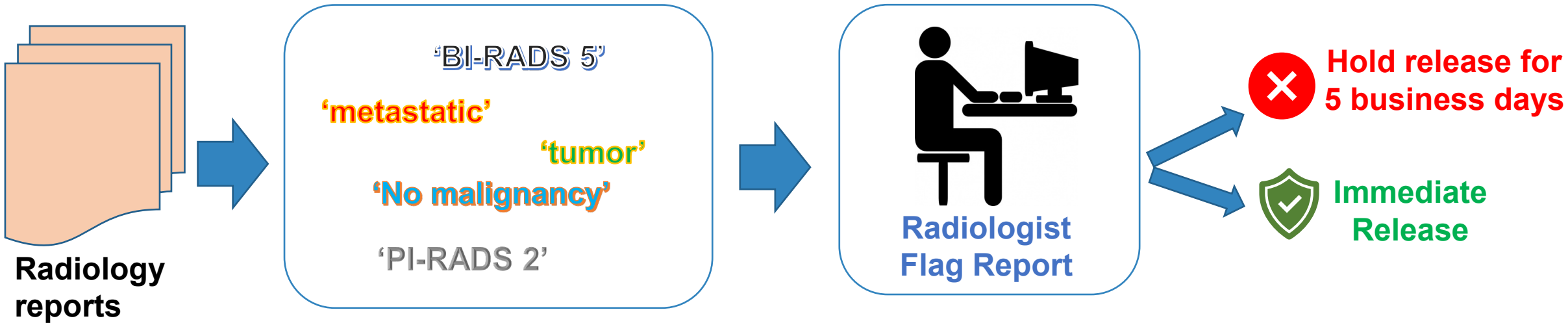
Goal

- Have a consistent, reliable automated flagging workflow that complies with both state and federal regulations

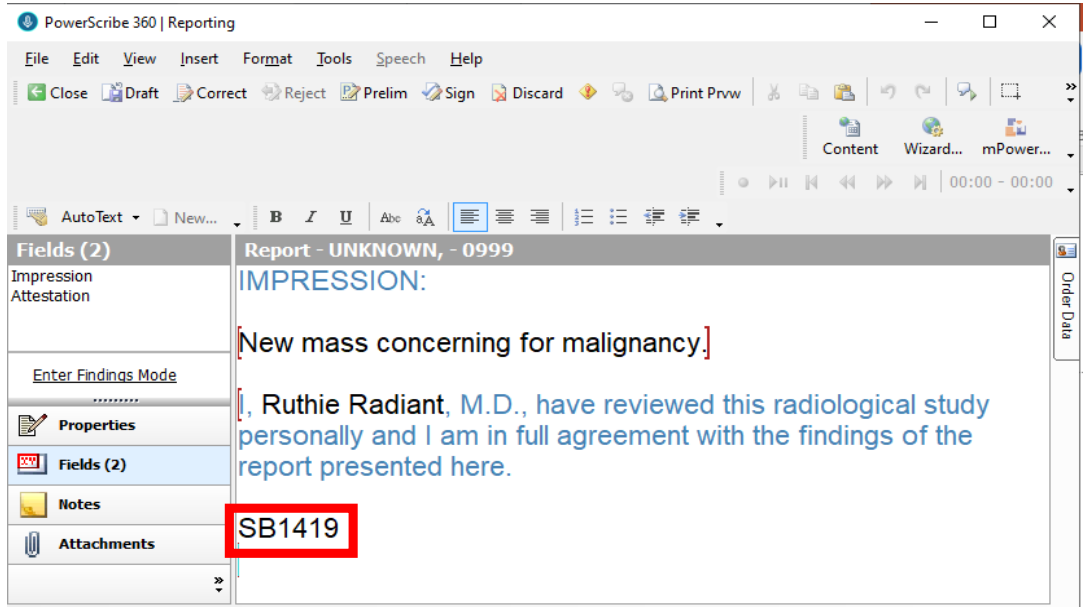
Solution

- Create an AI/NLP program to automatically analyze imaging reports and determine which to be held or released to patient portal

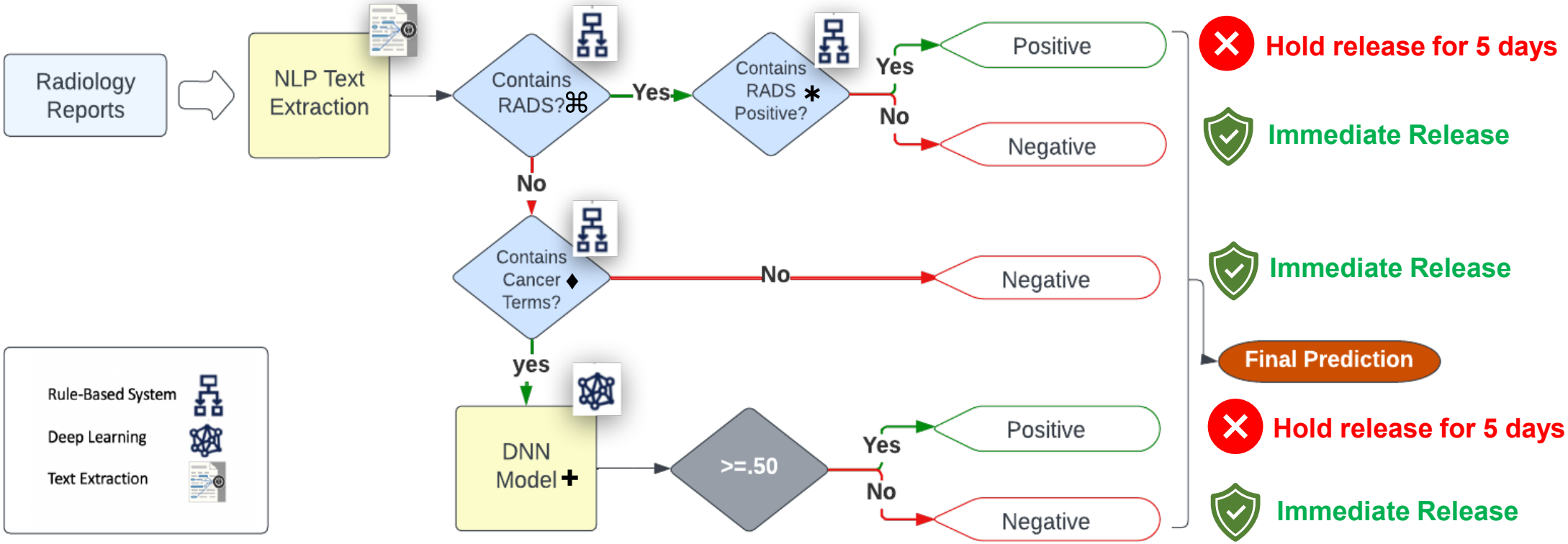
Manual Workflow: Overview



- Dictate phrase 'SB1419' at the bottom of the **IMPRESSION** section of the report.
- EHR will analyze for phrase dictated in reports from different reporting platforms (i.e. PowerScribe, CareConnect, and MagView)
- Manual flagging of reports to be held is onerous, prone to inconsistency, and difficult to standardize



Automated Workflow with AI: Overview



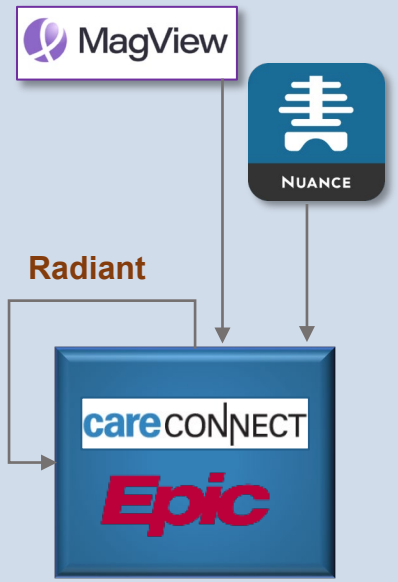
⌘ Pre-existing malignancy classification schemes (i.e. BI-RADS, LI-RADS)
✱ Positive defined as > 50% probability of malignancy on malignancy classification schemes (i.e. BI-RADS ≥ 4, LI-RADS ≥ 4)
♦ Pre-selected 12 most common keyword stems: Malignan, Cancer, Neoplas, Tumor, Carcinoma, Metastatic, Metastas, Lymphoma, Sarcoma, Leukemia, Melanoma, Lymphoproliferative
✚ Training dataset labeled by physician informaticists was used to train the deep neural network (DNN) model

Data Preparation for Hybrid Machine Learning Model

Pre-Modeling Phase

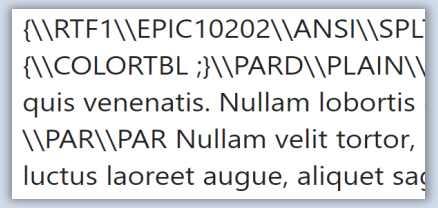
Deep Neural Network Modeling Phase

Data Collection

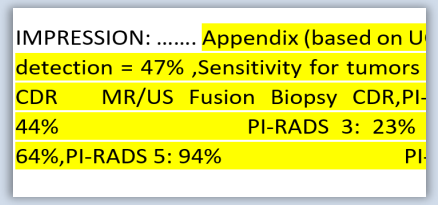


Radiology applications sending real-time data to EHR CareConnect

Data Cleaning

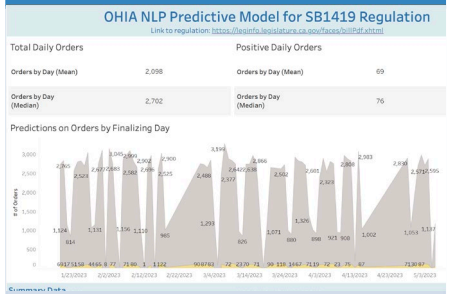


Handling rich text formats

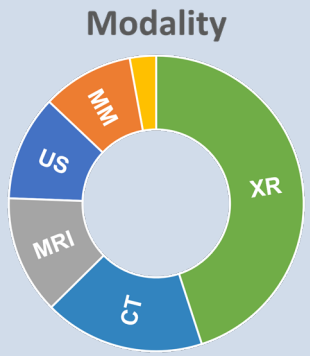


Removing appendixes

Exploratory Data Analysis



Daily volume of reports

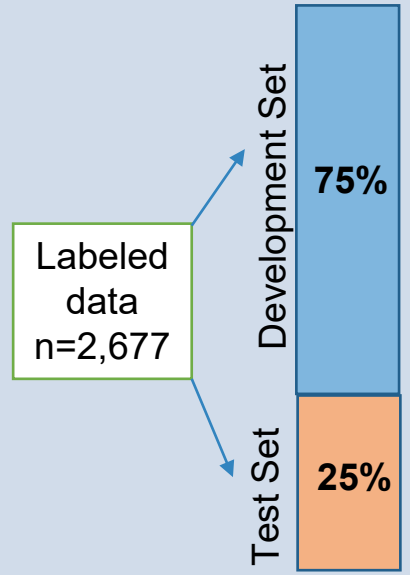


Distribution of imaging reports

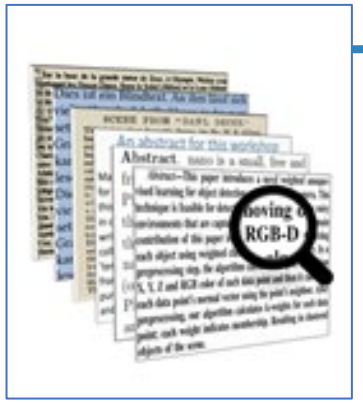
Data Engineering

- Normalize the text vocabulary
  - Case folding- reduce all letters to lowercase
  - Lemmatization- reduce inflected words to root words
  - Removing stop words
  - Tokenization using n-gram
- Term Frequency-Inverse Document Frequency: evaluates the frequency of words to determine its relevance in a given document

Data Splitting



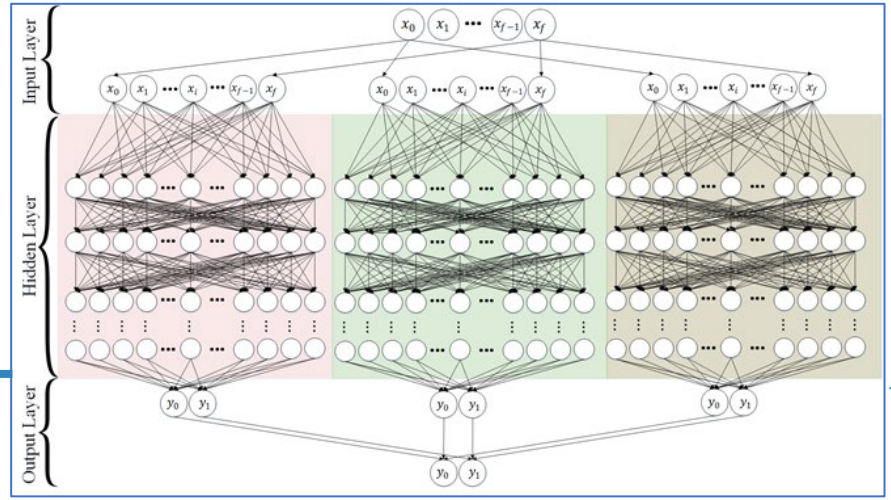
Model Development: NLP Deep Learning



Radiology Reports

NLP Data Engineering

Feature Extraction
(translate to Machine Language)

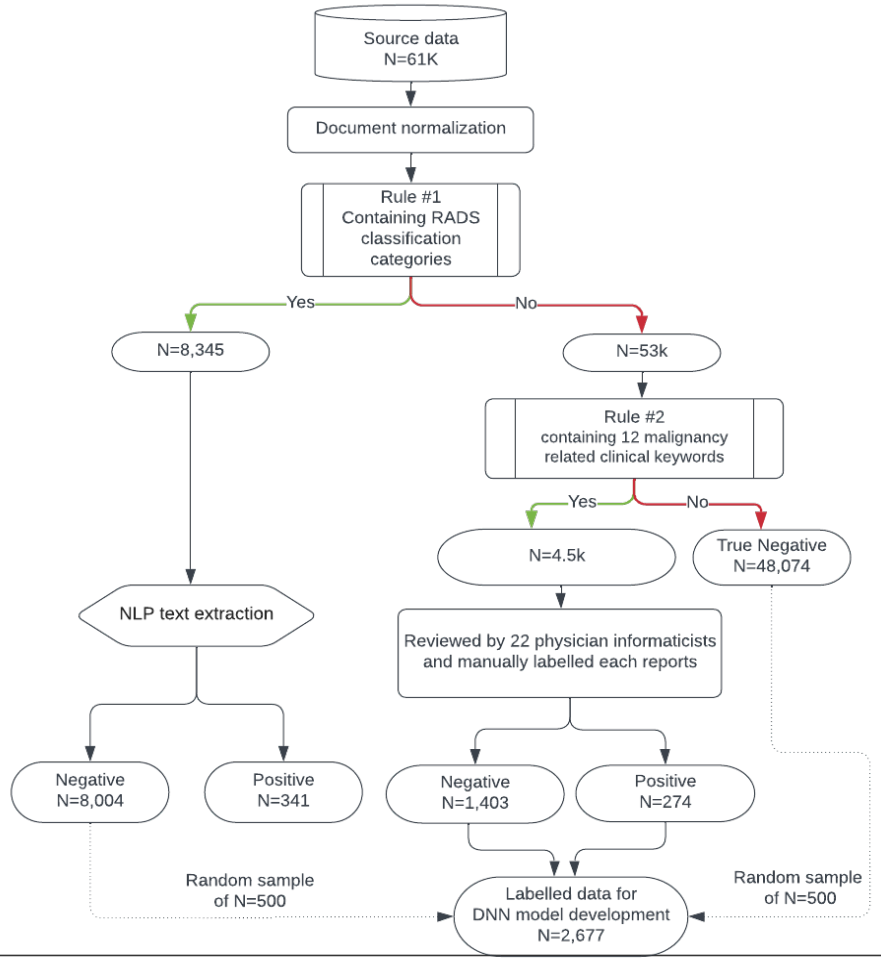


Deep Neural Networks' structure is designed to learn by multi-connection of layers (each layer only receives a connection from the previous and provides connections only to the next layer)

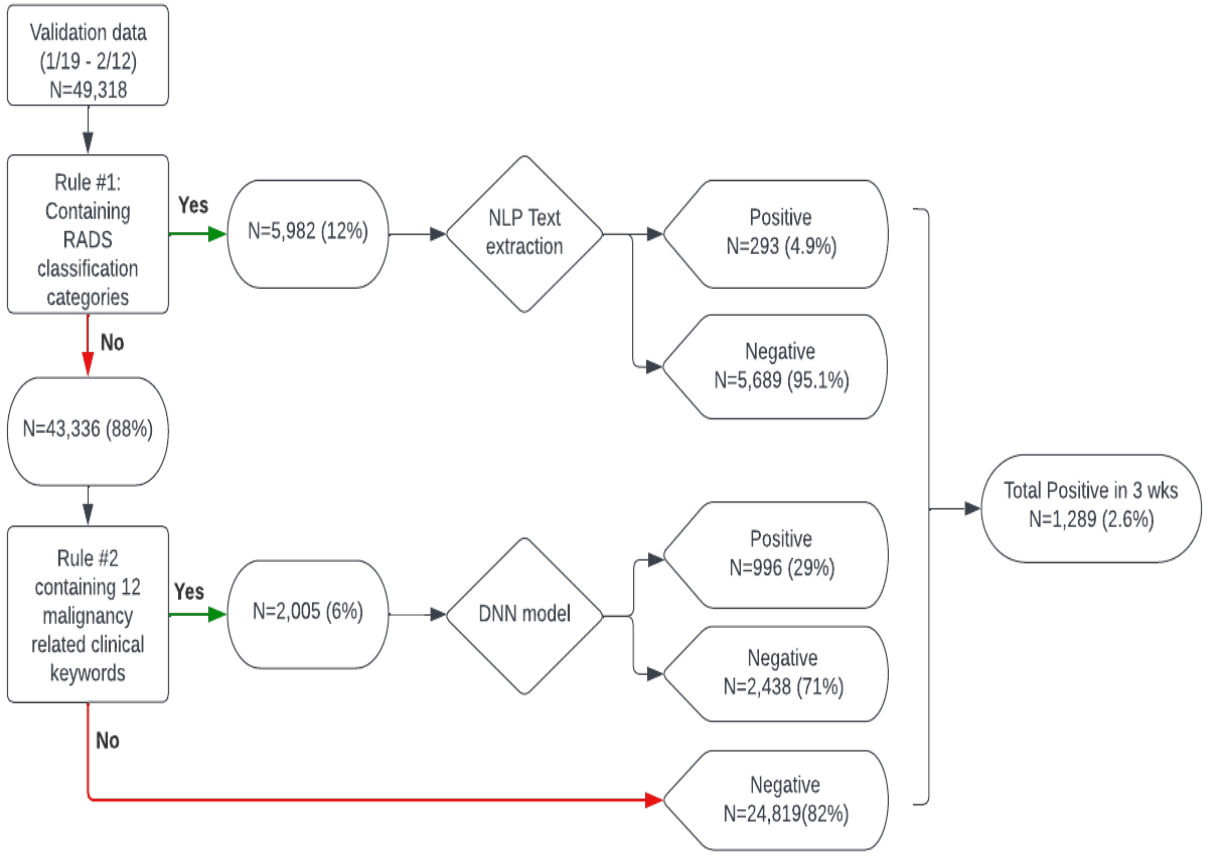
$Output$:
y_0 is used for not holding reports
y_1 is used to hold the radiology reports

Model Training and Validation

Prepare training dataset from reports (12/01/2022-12/31/2022)



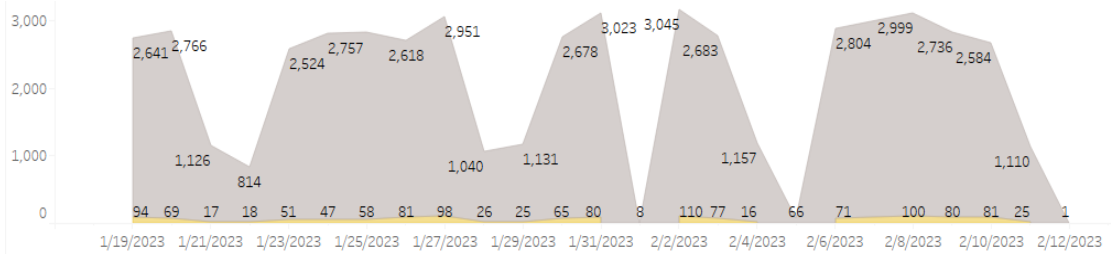
Apply AI model to reports (01/19/2023-02/12/2023)



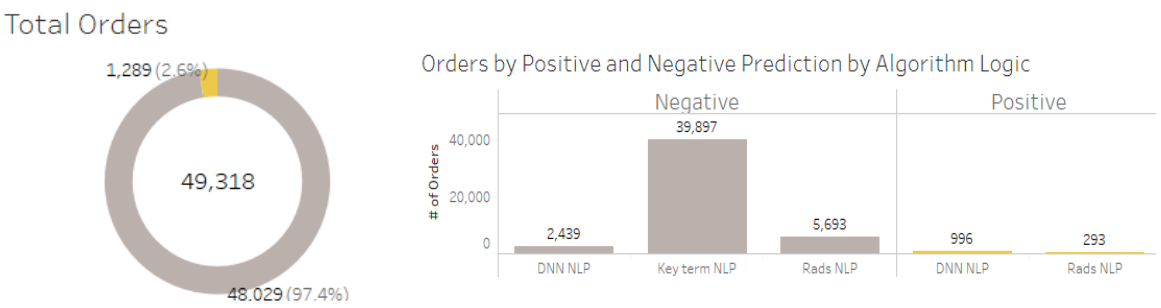
Model Evaluation

Applying AI Model

- Daily dashboard of AI model applied on reports that resulted from 01/19/2023-02/12/2023



- A total of 1,289 (2.6%) reports were positive and met the criteria for hold, with 996 reports identified from the DNN NLP and 293 reports from –RADS rule NLP, respectively.



Evaluating Performance

- Combining model output from both DNN and rule-base NLP
- Dataset: 25% of data from 12/01/2022-12/31/2022 (N=61K)

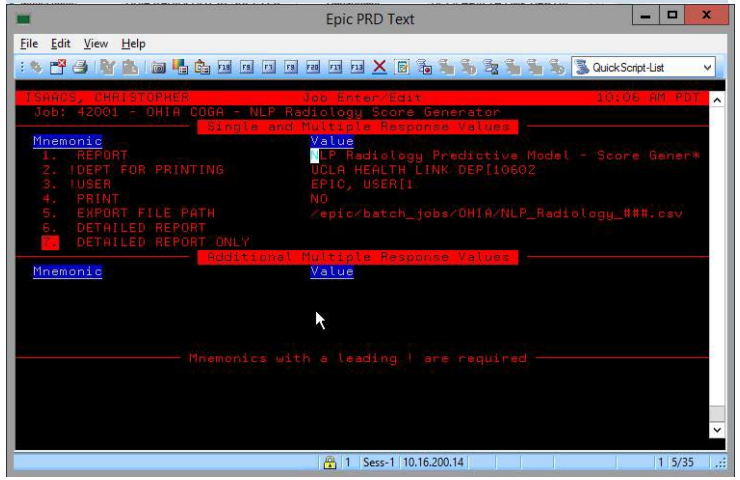
		Label Truth	
		True_Negative	True_Positive
Prediction	Negative	14535	27
	Positive	62	135

SE (Recall)	0.8336
SP	0.9958
PPV (Precision)	0.6857
NPV	0.9981
ACC	**0.9940**
F1-Score	0.7524

Model Monitoring

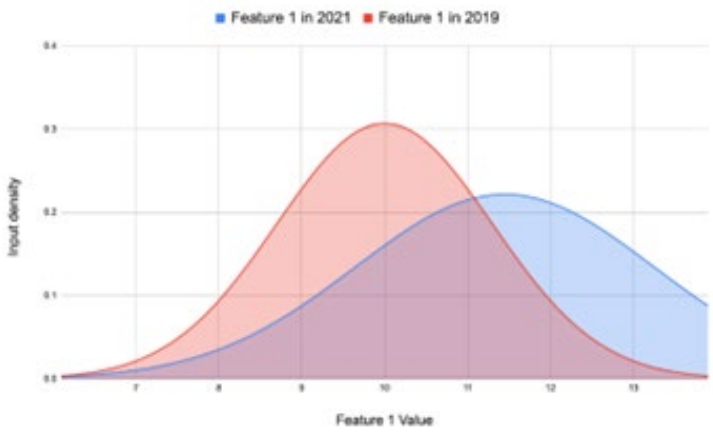
System Monitoring

- CareConnect build to notify users of system failure



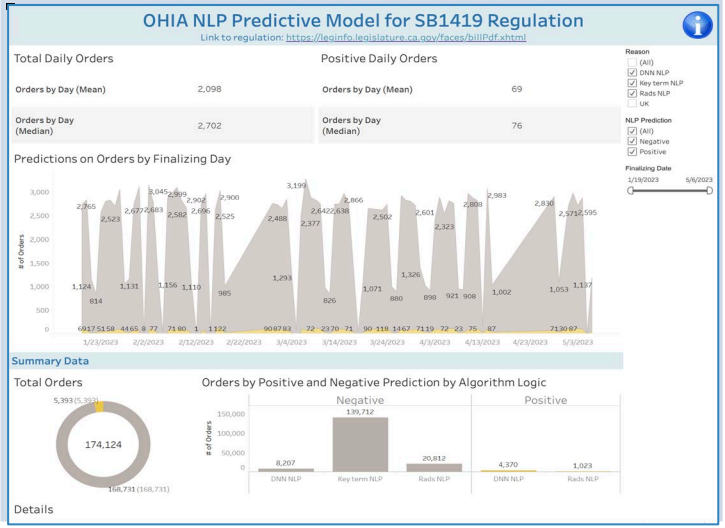
Performance Monitoring

- Periodically review of model performance post deployment
- Analyze for data drift or performance deterioration



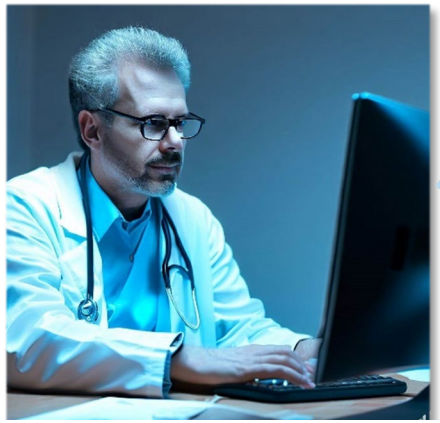
Business Monitoring

- Evaluate how the model impacts the business with dashboards



EHR Workflow Integration

- Awaiting further development by Epic EHR to incorporate AI model into Radiant workflow to perform hold functionality
- Status of released or held reports visible to providers



careCONNECT