Optimizing Best Practice Alerts Through Machine Learning

Population Health, CHIDS, MCIT, Predictive Analytics Unit (PAU)

Epic best practice alerts (BPA) are a mainstay of clinical decision support. They provide timely information to guide clinical care and ensure that the care team makes the best care decisions in a given clinical situation. Unfortunately, the number of alerts is huge, and, in some cases, nearly 99% of these alerts are ignored. In this project, we apply machine learning models to predict provider interaction with a best practice alert. Effective predictions can decrease the number of unnecessary alerts, and ultimately drive optimal care for our patients.

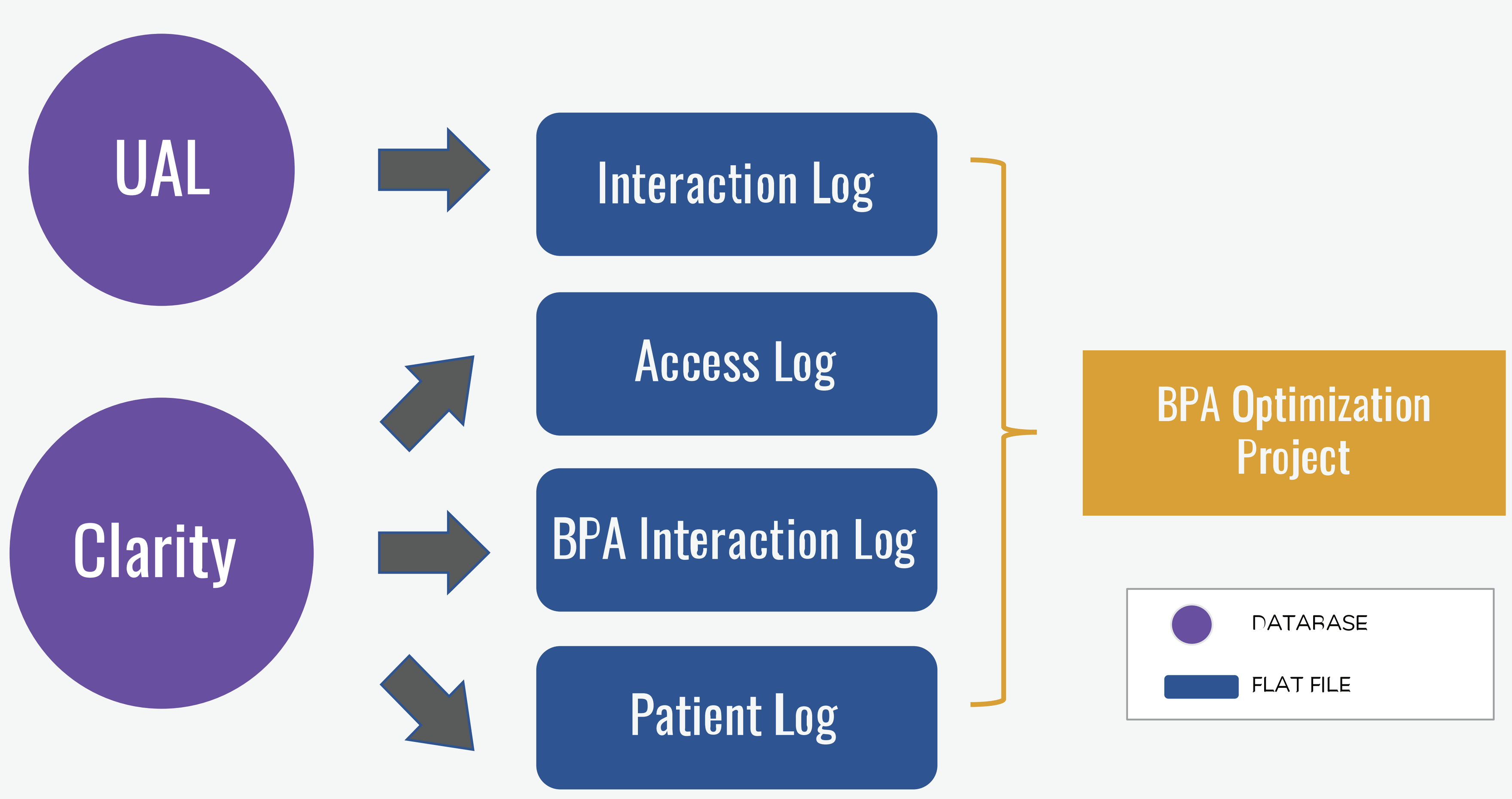


Figure 1: The BPA optimization projects uses data extracted from Epic. UAL and Clarity are databases stored in the Epic system. The Interaction Log stores activity records by Epic system users through clicking and typing; the Access Log contains supplementary information to the activity records in the Interaction Log; the BPA Interaction Log stores interaction records corresponding to individual BPAs; the Patient Log consists of each patient's latest demographics, and past encounters dated back to 2014. The Interaction Log, the Access Log and the BPA Interaction Log range from May to July, 2017.

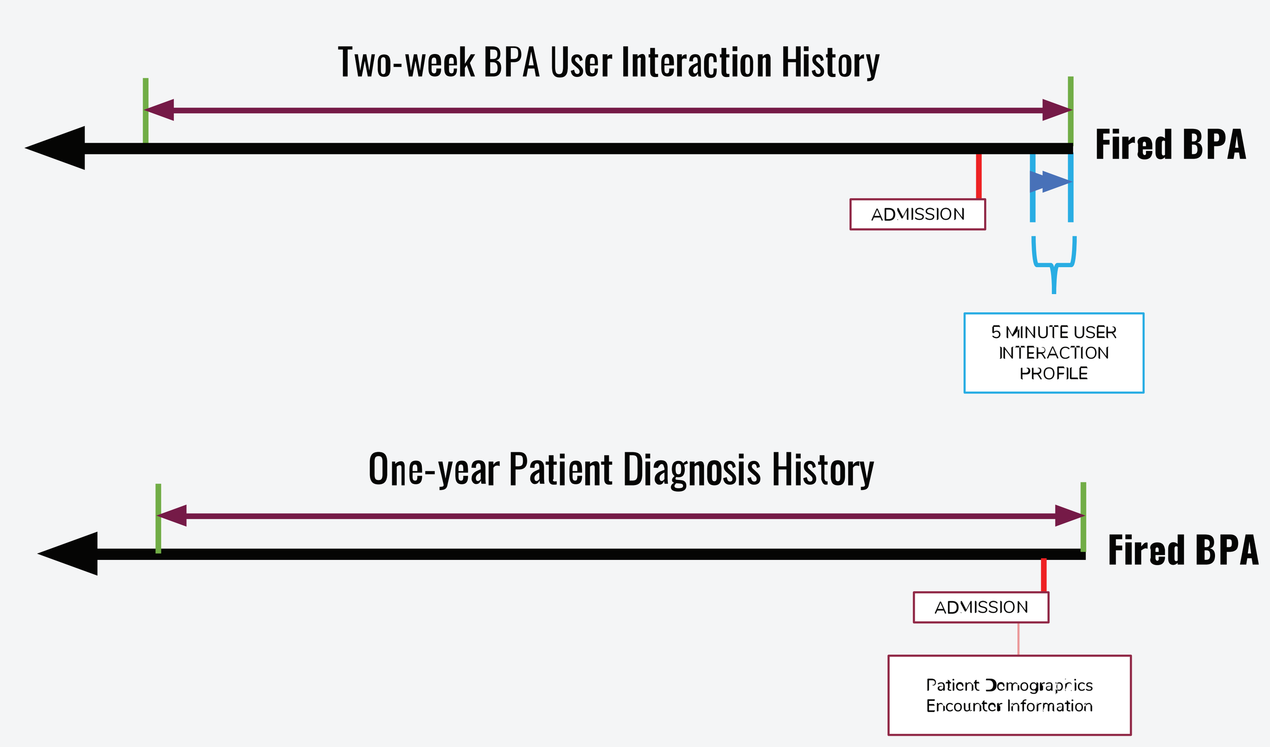


Figure 2: Experimental design: from the time a BPA is fired, user (clinician) information is extracted from two-week BPA interaction history, 5-minute user interaction profile, and provider types; patient information is extracted from 1-year diagnosis history, the patient's latest demographics and information related to the latest encounter.

PROJECT GOALS

Build machine learning models to predict interaction with a best practice alert, explore factors that drive best practice alert interaction and adoption, and implement interventions on these factors to attempt to drive more effective display and interaction with these clinical decision support tools.

CURRENT STATUS

We have built models to predict interaction with a pilot best practice alert for shingles vaccine. The model has good predictive power.

NEXT STEPS

Future work will identify causal influences and potential interventions to drive alert engagement.

OUR TEAM

The Predictive Analytics Unit aims to drive the successful translation of prediction models into clinical practice. We apply the latest advances in machine learning and predictive analytics toward translating prediction and classification problems across the medical center.

Please Email Us with Questions predictiveanalytics@nyumc.org

Deep Learning-Based Extraction of Suspect Diagnoses from Historical Outpatient Visits for Real-Time Recommendation

Population Health, CHIDS, MCIT, Predictive Analytics Unit (PAU)

In this project, we work closely with MCIT Clinical Informatics and clinical leaders to build deep learning based models to recommend HCC (Hierarchical Condition Category) diagnoses to update at a patient's outpatient visit. Optimal identification and addressing of chronic diagnoses from year to year ensure that our patients receive the best care for all of their morbidities.

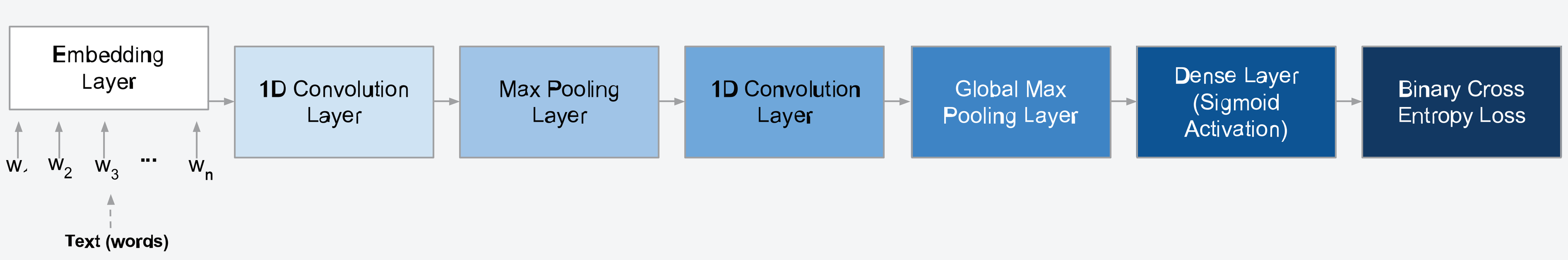


Figure 1: Preliminary convolutional neural network architecture used for building classification models. Input to models is raw clinical text from outpatient notes; output is classification label corresponding to diagnosis code(s).

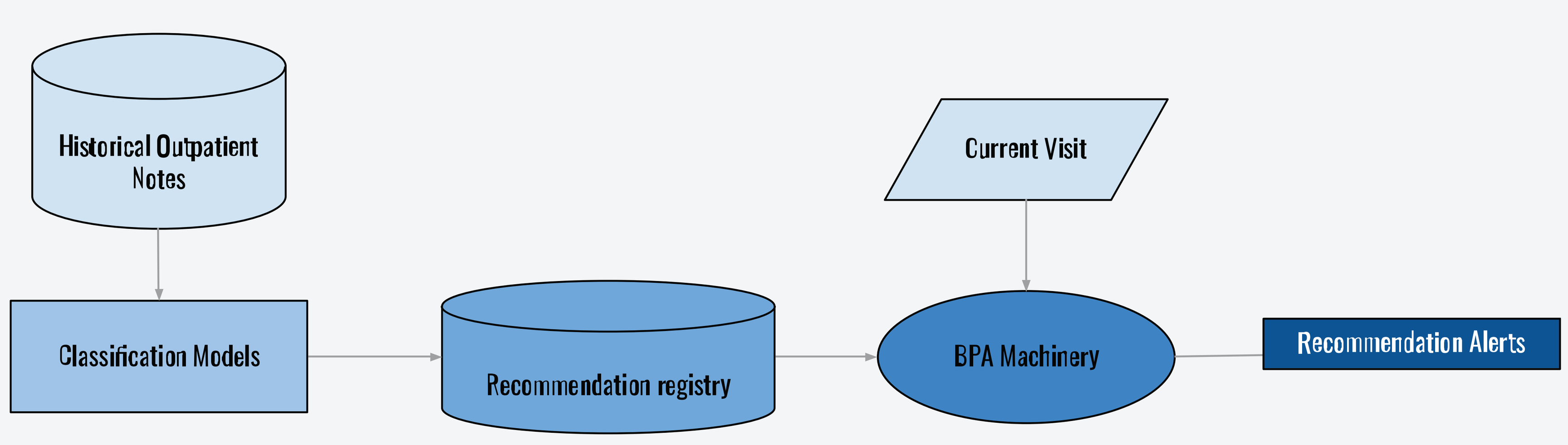
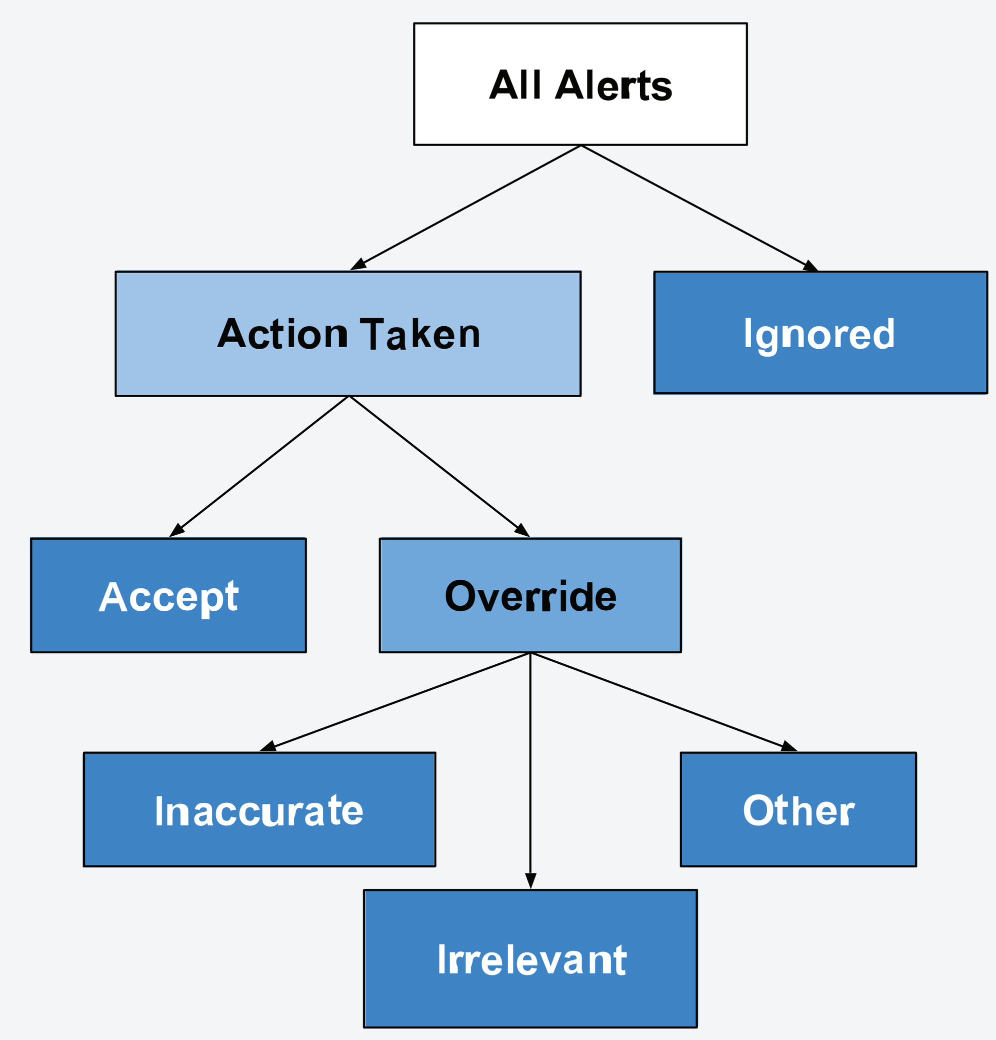


Figure 2: Overall operational workflow - Historical notes from Epic Clarity are processed through classification models to assign suspect diagnoses, which are added to the recommendation registry database. During the current visit for a patient, BPA machinery fires suspected diagnosis alerts pre-computed in registry. Based on provider action, registry is updated.



$$Precision = \sum \frac{\frac{Accepted}{Acceted+Override}}{Total\ Alert\ Count}$$

Best precision is 1.0: All accepted, no overrides or ignores
Worst precision = 0.0: All overrides or ignores, no accepts

Figure 3: End-user alert effectiveness. BPA machinery allows different actions for providers in response to diagnosis recommendation alerts. Actions include: a) Accept: Add diagnosis; b) Override: override alert with a reason - inaccurate diagnosis, irrelevant to practice of current visit or other (non-specific catch-all); c) Ignore: BPA is ignored without any action. Information retrieval-based metrics, such as precision, provide a way to evaluate alert effectiveness. High precision is achieved if all alerts displayed result in "Accept" user action.

PROJECT GOALS

Apply deep learning to 5.5 million outpatient notes to classify patients according to diagnoses that they should have, validate through human review the effectiveness of these algorithms for identifying diagnoses, provide effective delivery of this information to providers through Epic's suite of clinical interventions and tools, demonstrate a more robust capture of patient morbidities through machine learning based classifiers and effective clinical decision support.

CURRENT STATUS

We have selected 10 high-yield HCC groups to pilot our recommendation infrastructure - diabetes, vascular disease, arrhythmias, morbid obesity, CHF, COPD, angina, arthritis, cancers and depression.

NEXT STEPS

To build deep learning models to assign 111 ICD codes from these groups to outpatient visit notes. We are simultaneously (a) designing information retrieval based methods to evaluate effectiveness of our recommendations, (b) moving toward clinical validation.

OUR TEAM

The Predictive Analytics Unit aims to drive the successful translation of prediction models into clinical practice. We apply the latest advances in machine learning and predictive analytics toward translating prediction and classification problems across the medical center.

Please Email Us with Questions predictiveanalytics@nyumc.org

Enterprise Predictive Analytics Infrastructure at NYULH

Population Health, CHIDS, MCIT, Predictive Analytics Unit (PAU)

With the advent of ubiquitous big data, medical predictive analytics have taken center stage toward facilitating real-time clinical decision-making, increasing operational efficiency, reducing costs, improving quality of care, and providing the foundation for a learning healthcare system. In this work , we describe our design of a predictive analytics Infrastructure that meets the goals of a production level deployment as opposed to a scientific project. Our implementation takes advantage of modern technologies and best practices in systems engineering to produce a modelling infrastructure that is reliable, scalable, and easy to maintain and deploy.

PAU Operational/ Data Flow

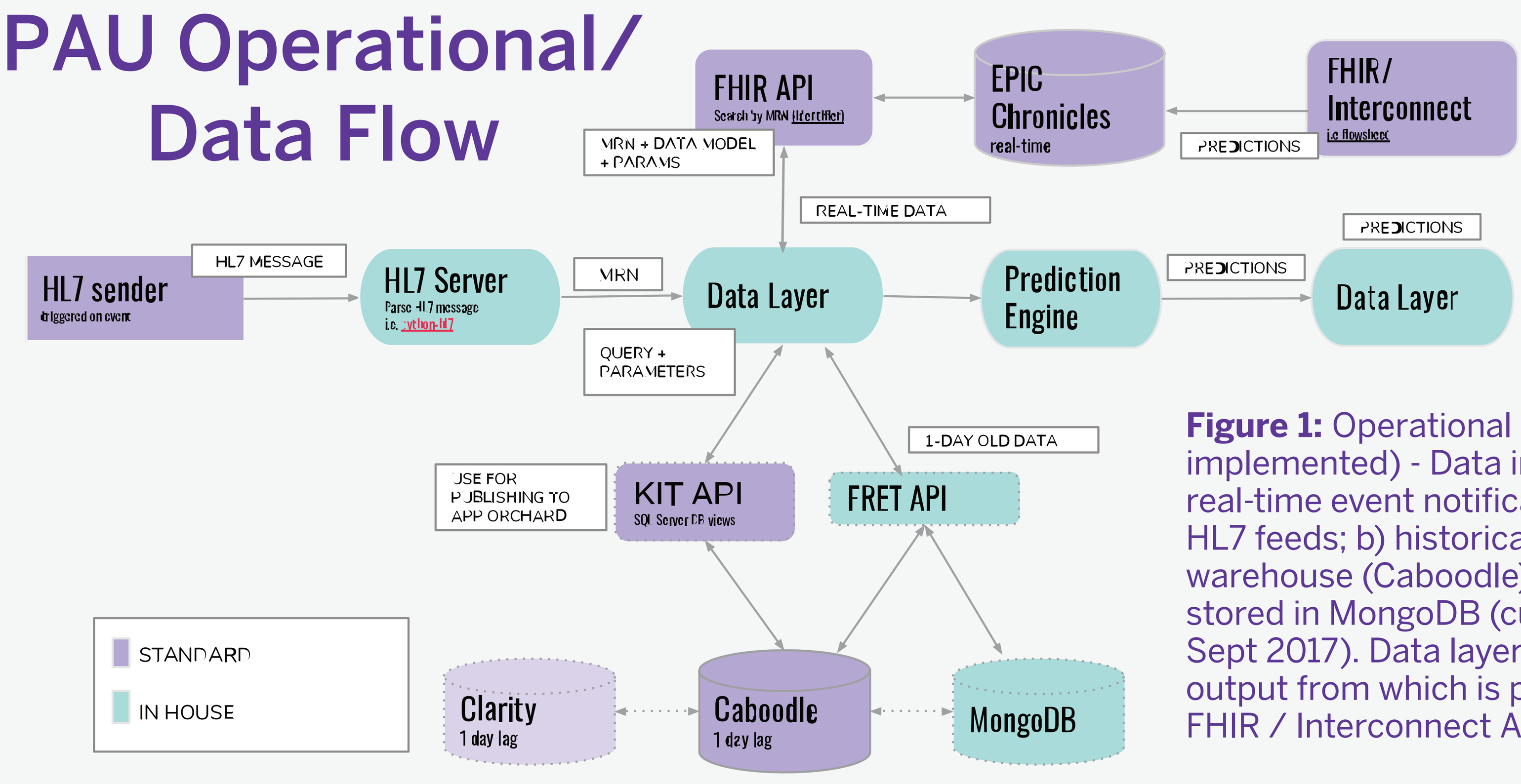


Figure 1: Operational data workflow (partially implemented) - Data infrastructure consumes a) real-time event notifications directly from Epic via HL7 feeds; b) historical data from Epic data warehouse (Caboodle) or Epic Clarity snapshot stored in MongoDB (currently has data from 2014-Sept 2017). Data layer feeds the prediction engine, output from which is populated back into Epic via FHIR / Interconnect API.

Full Pipeline

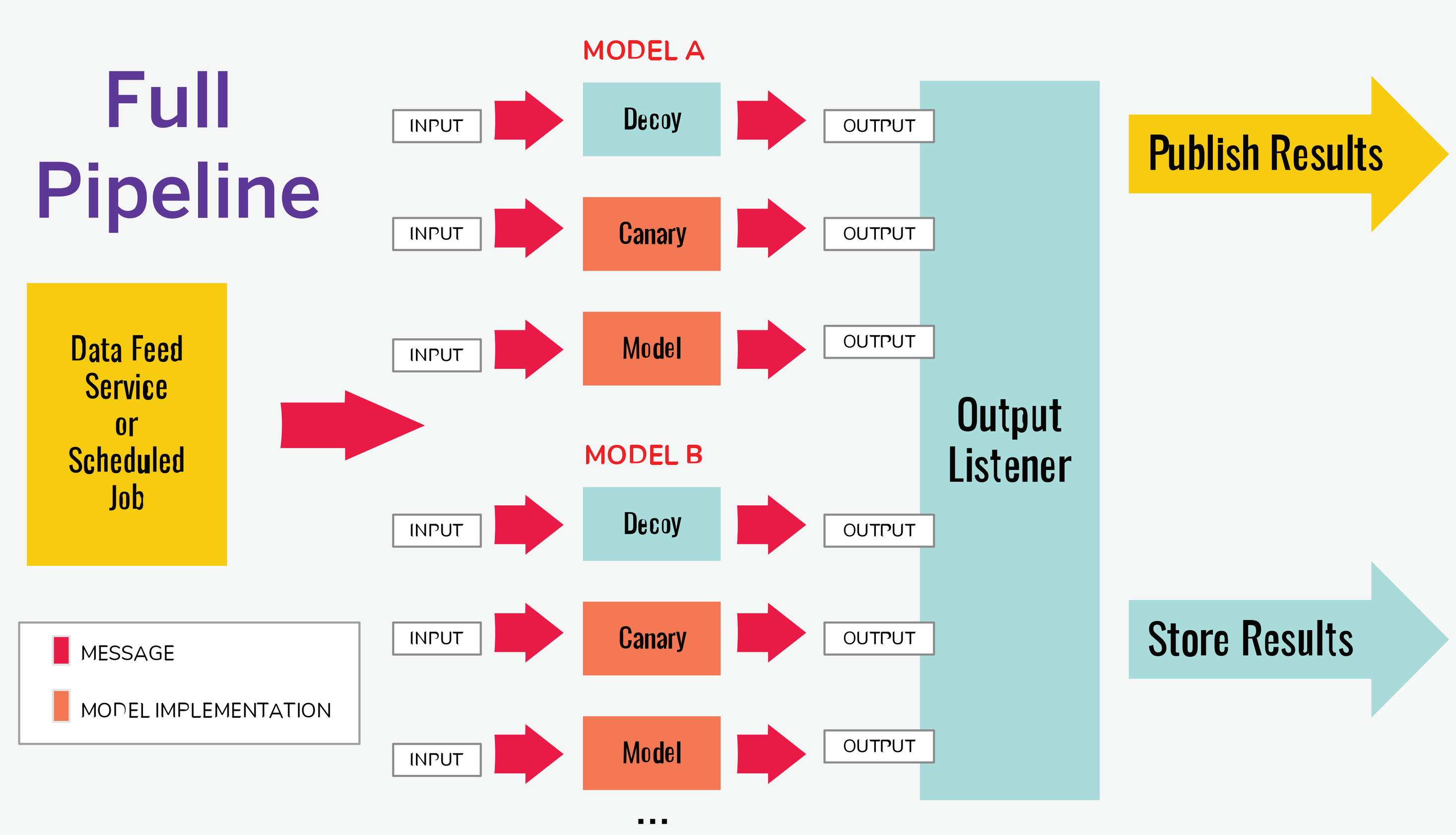


Figure 2: Production Model Deployment Architecture. Models implement a standardized interface and are deployed as containerized services that: a) consume input data stream; b) publish results for downstream services to listen and act on. Production models can be supplemented with Decoys (to capture raw input for later re-use) and Canaries (to monitor drift in model or data). Streaming also facilitates comparing new test models that consume same input stream as production model.

Machilne Learning Environment

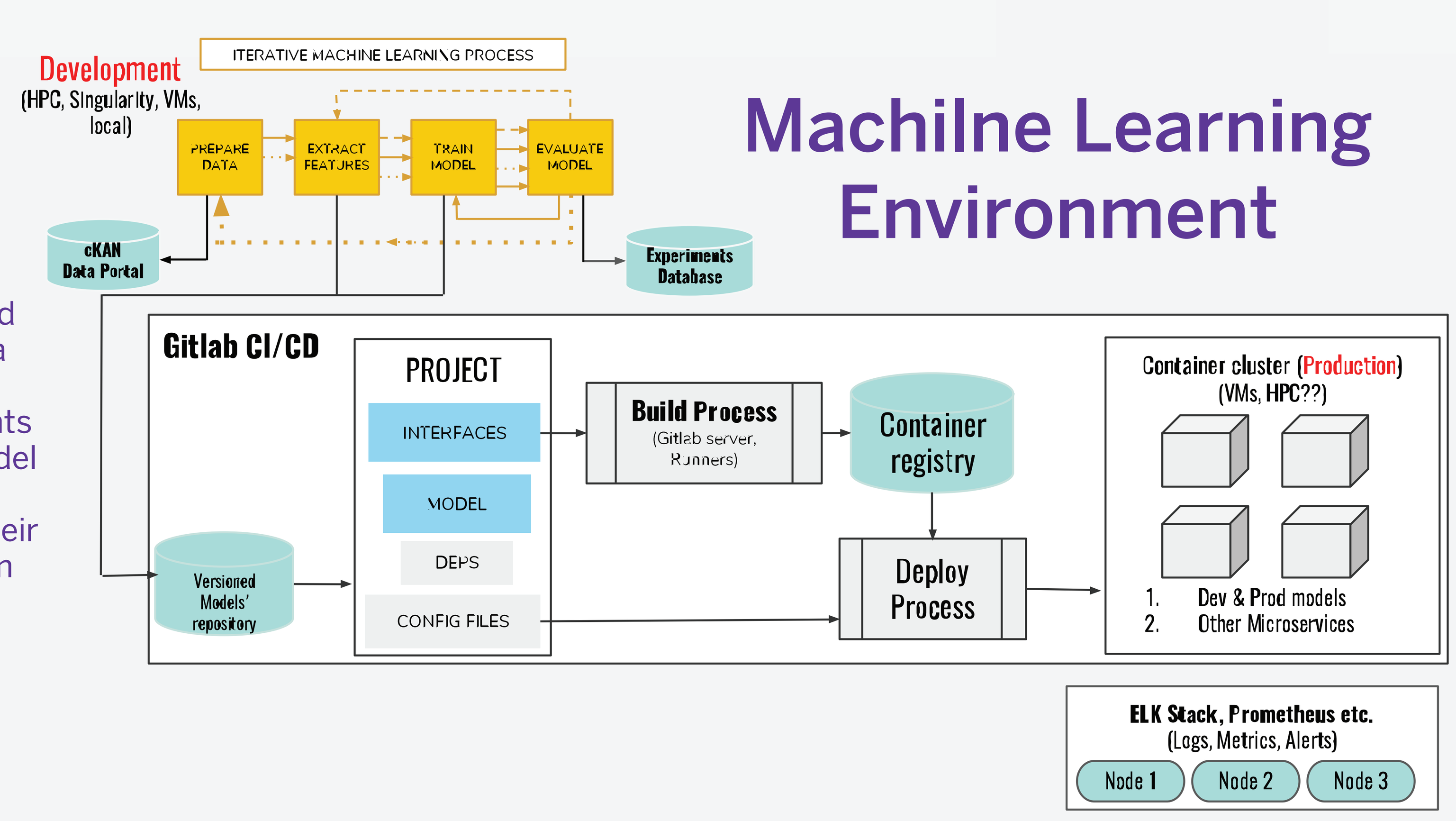


Figure 2: The Vision - A Dev cycle using High Performance Computing (HPC), Singularity, Virtual Machines (VMs) and local resources to (1) build datasets and store them (cKAN data portal), (2) build machine learning models, and (3) store the experiments (experiments database). Once a model is built, the Gitlab CI/ CD allows versioning models, containerizing their deployment and finally pushing them into a Prod Container cluster for scalable deployment through microservices. Finally logs and monitoring are captured using ELK stack and Prometheus.

PROJECT GOALS

Our goal is to come up with a centralized infrastructure that, once in place, would support real-time data extractions, make the deployment of models from development to production seamless and require minimum engineering input from data scientist and external teams.

CURRENT STATUS

We have successfully completed the deployment of this enterprise prediction infrastructure using a published predictive model for Congestive Heart Failure prediction from electronic health record. The main components of this project have been fully implemented since March 2017 and have been running in production environment for over a year.

NEXT STEPS

This production-level infrastructure served as a proof of concept of our approach. It has been successfully reused for our second production project, 2-month Mortality Prediction, and has also been generalized to multiple projects currently in development.

OUR TEAM

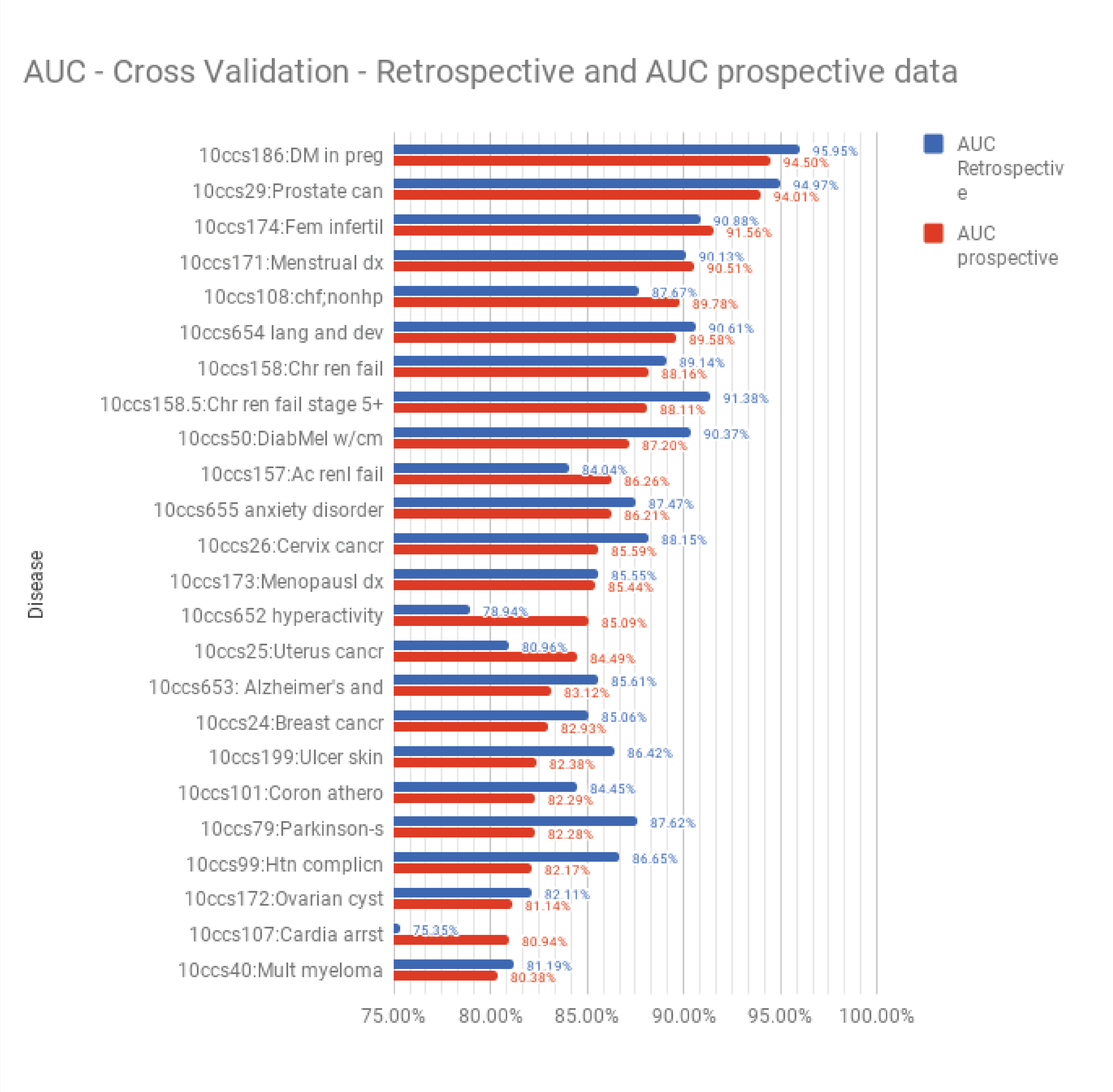
The Predictive Analytics Unit aims to drive the successful translation of prediction models into clinical practice. We apply the latest advances in machine learning and predictive analytics toward translating prediction and classification problems across the medical center.

Please Email Us with Questions predictiveanalytics@nyumc.org

Efficient Construction of Machine Learning Based Models for Forecasting and Prediction

Population Health, CHIDS, MCIT, Predictive Analytics Unit (PAU)

Building machine learning models from electronic health records typically involves arduous data collection, conversion of the data to machine readable formats, and applications of machine learning algorithms to this data. This process can take months of time. Recent work has demonstrated generalized systems to build models off electronic health record data. In this project, we have built a generalized infrastructure that ingests electronic health record data and builds machine learning models for any outcome of interest in the span of hours rather than days. This includes a rigorous framework for daily validation of the models, via prospective and retrospective validations and final validation through human review.



Top 20 Disease families with highest forecasting area under ROC curve, within 1 year.

PROJECT GOALS

Building infrastructure and pipelines that allow the specification of arbitrary outcomes, inclusion criteria, exclusion criteria, type of the machine learning model (including deep learning), and type of validation and analysis. The system has the following features: (1) Fast one-click model building with extensive evaluation within minutes to hours for any task of interest. (2) Application of best practices and the state of the art ultilization of information in diagnoses, procedures, encounters, medications, labs, vitals, and clinical notes in the tasks of predicting any outcomes of interest.

CURRENT STATUS

We have demonstrated building models for 200 prediction and disease forecasting tasks in the span of hours versus days or months. All results have been evaluated on retrospective and prospective cohorts. Top predictable diseases are included in Figure 1.

NEXT STEPS

We are working with clinical partners to identify opportunities for interventions in renal failure and breast cancer.

OUR TEAM

The Predictive Analytics Unit aims to drive the successful translation of prediction models into clinical practice. We apply the latest advances in machine learning and predictive analytics toward translating prediction and classification problems across the medical center.

Please Email Us with Questions predictiveanalytics@nyumc.org

Optimizing Patient Care at the End-of-Life with Machine Learning

Population Health, CHIDS, MCIT, Predictive Analytics Unit (PAU)

In this project, we work closely with MCIT Clinical Informatics, and the Value Based Medicine (VBM) clinical task force to build machine learning models to deliver timely supportive care interventions to patients approaching the end of their life. Our goal is to ensure that terminally ill patients have the time and means to discuss and define their goals of care and end of life wishes to ensure they receive timely and appropriate treatments that align with their wishes.

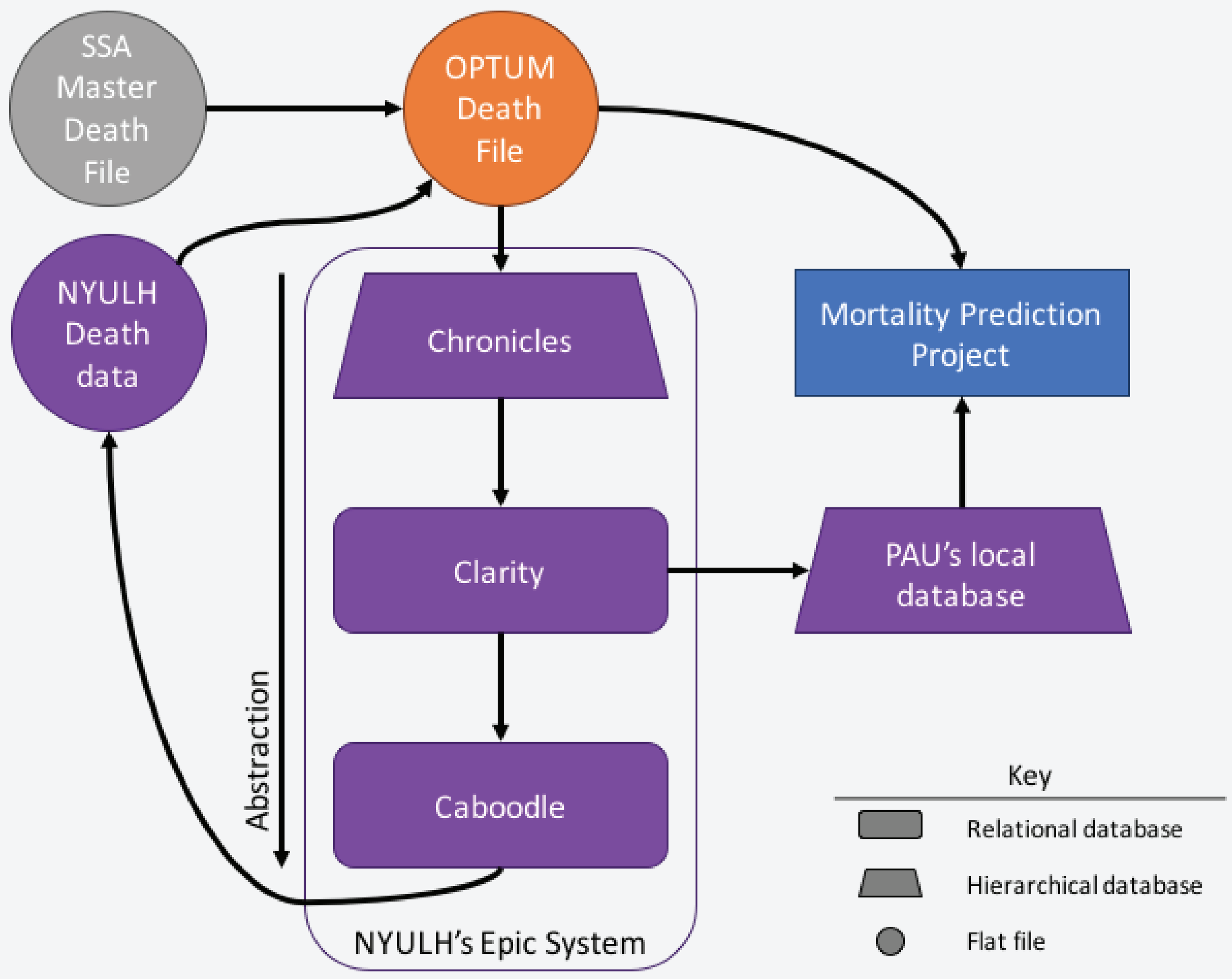


Figure 1: The end-of-life project uses data extracted from Epic and supplements the outcomes with data derived from the Social Security Administration.

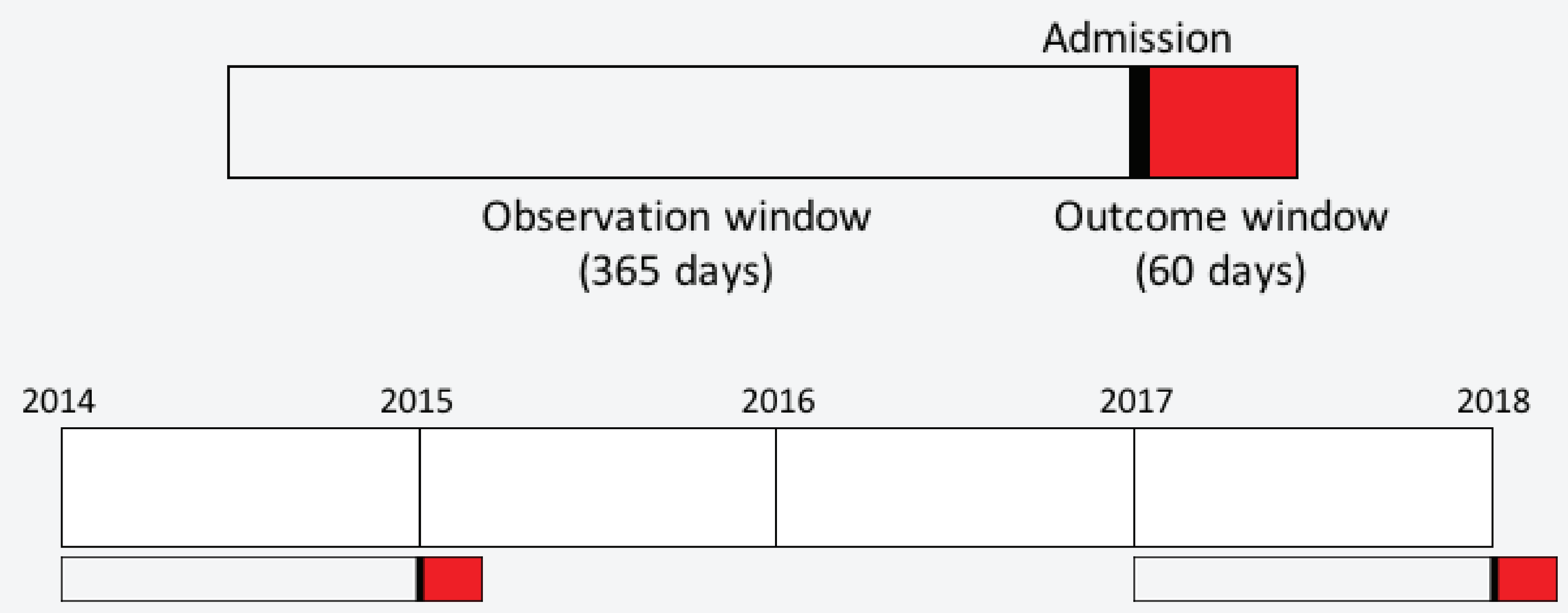


Figure 2: Experimental design. Top: from the time of an inpatient admission, the previous 365 days of data is considered, i.e. nothing from the day of admission, to predict patient outcomes 60 days into the future. Bottom: with four years of data and requiring 365 days of an observation window, three complete years of inpatient admissions are included.

PROJECT GOALS

Ingest electronic health record data prior to a patient's visits to apply the latest in machine learning to determine a patient's risk of death. Provide timely delivery of risk estimates to point of care provider teams to incorporate into their clinical care. Demonstrate how machine learning based systems working with human expert providers can yield more appropriate decisions in a timely manner to directly influence patient care.

NEXT STEPS

Future work will further incorporate additional data types including labs, medications, vitals, and clinical notes and apply deep learning techniques for improved modeling, as well as integration with Epic to facilitate real-time prediction and notification to improve patient care.

CURRENT STATUS

We have built machine learning models from clinical data that include claims, procedures, and encounters to predict risk. An automated email is generated daily and sent to select providers for human feedback and review. Results have been mostly positive with providers able to deliver care interventions in a timely and effective manner.

OUR TEAM

The Predictive Analytics Unit aims to drive the successful translation of prediction models into clinical practice. We apply the latest advances in machine learning and predictive analytics toward translating prediction and classification problems across the medical center.

Please Email Us with Questions predictiveanalytics@nyumc.org