Georgia Tech Students' work building PLP UI

ABHISHEK KHOWALA

Landscape

- CS 6440 – Introduction to Health Informatics course give students opportunity to work on real world high impact healthcare related projects in team supported by External Mentor and TA Mentor
- ~15 weeks long project
- Typical team size of 4-6 students
- The delivery of project is broken into 6 milestones
- Infrastructure for continuous build and deployment provided
- FHIR servers provided

Deliverables

- D0 – Team formation & Team Confirmation (WW03)
- D1 – Project Topic: Team Topic, Research, Use Cases, Project Plan (WW06)
- D2 – Technical Presentation: Technical Architecture (WW09)
- D3 – Team Progress: Effectively show elements of application (WW12)
- D4 – Full demonstration of the application (WW15)
- D5 - Final Project Submission (WW15)

- Students provide mentors with weekly update irrespective of any major delivery or not
- Each deliverable includes presentation as well as a narrated video link, every student in the team need to participate

DEPLOYMENT INFRASTRUCTURE

FOR CONTINUOUS BUILD, CONTAINERIZATION AND DEPLOYMENT INTO A SERVER FARM

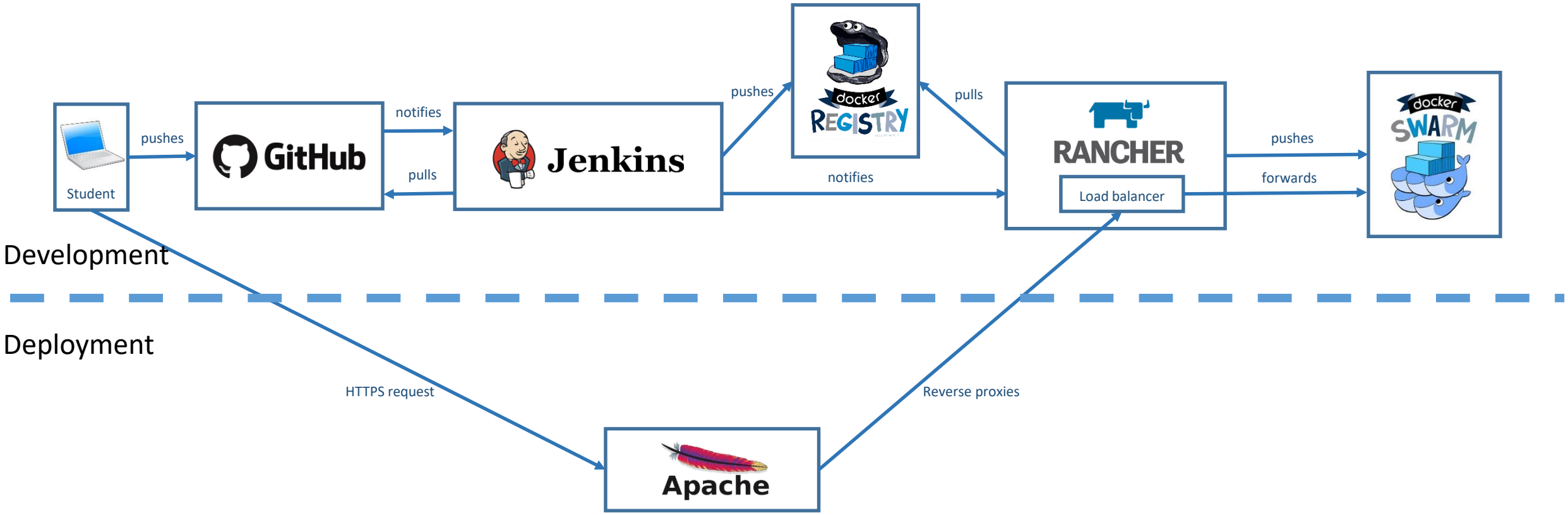
HDAP – Health Data Analytics Platform

- Software platform designed to support students in conducting health oriented projects at Georgia Tech
- HDAP role in projects
  - Access to FHIR servers
  - Application Deployment
- HDAP 3 core functions
  - Provide synthetic and de-identified sources of healthcare data
  - Provide tools to analyze these data
  - Provide a healthcare application development and hosting environment and app galleries

HDAP – Continued:

- FHIR SERVER ACCESS PROVIDED FOR
  - **MIMIC (STU3)**
    - OMOP data
    - **Endpoint:** https://apps.hdap.gatech.edu/gt-fhir/tester/
  - **SyntheticMass (STU3)**
    - Endpoint: https://apps.hdap.gatech.edu/syntheticmass/
  - **HAPI (R4)**
    - Endpoint: https://apps.hdap.gatech.edu/hapiR4/

Development/Deployment Pipeline



Process

- Student pushes code to GT GitHub
- GT GitHub Notifies Jenkins of push
- Jenkins builds project and pushes images to the private Docker registry
- Jenkins notifies Rancher of the push to the registry
- Rancher pulls the images, builds containers, and deploys them to the Docker Swarm

OHDSI PLP UI Project

FALL 2018 SEMESTER:
CONSUMER FACING APPLICATION FOR PATIENT LEVEL PREDICTION

TEAM COMPOSITION

TEAM MEMBERS

MICHAEL GARNER

JONATHAN FUNG

ALYSSA DE LEON

JACOB GILBERT

CHUREN (CHU) SHAO

TA MENTOR

ABHISHEK KHOWALA

EXTERNAL MENTOR

DR. PATRICK RYAN

DR. DAVID MADIGAN

HEAD TA

TIA POPE

PROFESSOR

DR. JON DUKE

PROJECT GOALS

- Analyze 3 example OHDSI PLP Models viz. AfibStroke, DepressionSuicide, NSAIDGIBleed
- Create a single interface that takes any PLP results object (such as 3 examples provided) and serve up an ordered set of questions to allow a target population user to estimate a personalized prediction of their risk of the outcome
- Use coefficient and covariate summary together with user answers to yes/no questions to produce the predicted probability and estimated confidence interval (CI) around those predictions.

Project Planning: Gantt Chart

Jumping Dog Runners

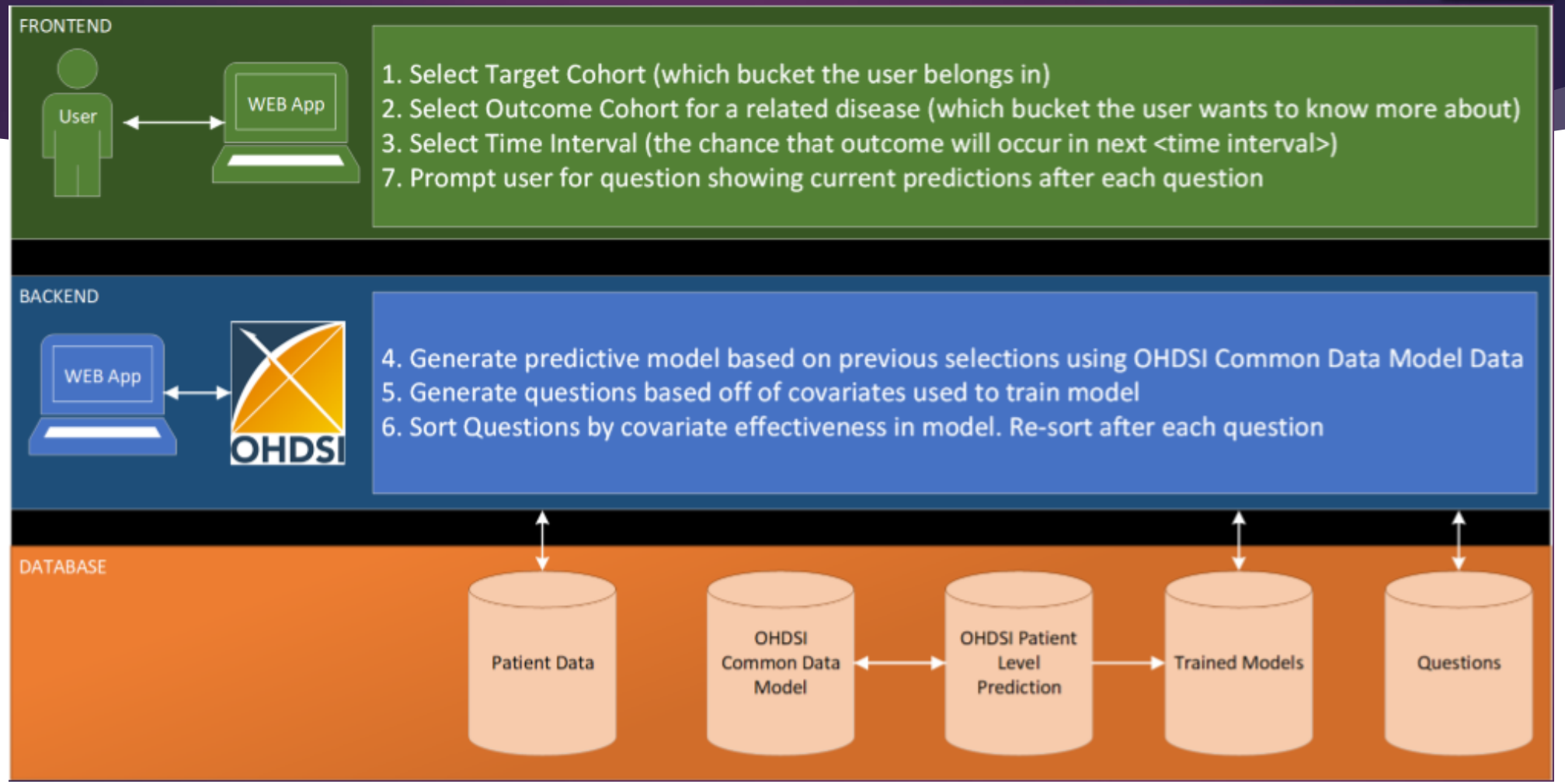
PROJECT TITLE	CS6440 Team Project
PROJECT MANAGER	Churen Shao

TASK TITLE	TASK OWNER	START DATE	DUE DATE	PCT OF TASK COMPLETE
Deliverable - 1 Team Topic Presitation				
Gantt Chart	Michael Garner	10/1/18	10/7/18	100%
Research	Jacob Gilbert	10/1/18	10/7/18	100%
Business Cases	Alyssa De Leon	10/1/18	10/7/18	100%
Introduction & Conclusion	Jonathan Fung	10/1/18	10/7/18	100%
Narration	Churen Shao	10/1/18	10/7/18	100%
Submission	Churen Shao	10/1/18	10/7/18	100%
Deliverable 2 - Technical Presentation				
Research	All	10/8/18	10/21/18	100%
Architectural Diagram	Jonathan and Jacob	10/8/18	10/21/18	100%
Presentation	Alyssa and Churen	10/8/18	10/21/18	100%
Narration	Jonathan and Alyssa	10/8/18	10/21/18	100%
Gannt Chart Update	Michael and Churen	10/8/18	10/21/18	100%
Meeting with Mentors	All	10/8/18	10/21/18	100%
Submission	Churen Shao	10/8/18	10/21/18	100%
Deliverable 3 - Progress Report				
Research (Questions and Model)	All	10/22/18	11/11/18	100%
Setup Development Environment	Churen Shao	10/15/18	11/4/18	100%
Front-End Development	Jonathan and Alyssa	10/22/18	11/11/18	100%
Back-End Development	Michael and Churen	10/22/18	11/11/18	100%
Testing and Q.A.	Jonathan and Jacob	10/29/18	11/11/18	100%
Progress Presentation	Alyssa and Churen	11/5/18	11/11/18	100%
Gannt Chart Update	Michael and Churen	10/22/18	11/11/18	100%
Narration	Michael	11/8/18	11/11/18	100%
Meeting with Mentors	All	10/22/18	11/11/18	100%
Submission	Churen Shao	10/22/18	11/11/18	100%
Deliverables 4 and 5				
Front-End Development	Jonathan and Alyssa	11/12/18	12/02/18	100%
Back-End Development	Michael and Churen	11/12/18	12/02/18	100%
Testing and Q.A.	Jonathan and Jacob	11/12/18	12/02/18	100%
Manual	Jacob and Churen	11/12/18	12/02/18	100%
Special Instructions	Jonathan	11/12/18	12/02/18	100%
Presentation	Alyssa and Churen	11/12/18	12/02/18	100%
Narration	Jacob	11/12/18	12/02/18	100%
Gannt Chart Update	Michael and Churen	11/12/18	12/02/18	100%
Meeting with Mentors	All	11/12/18	12/02/18	100%
Delivery	Churen Shao	11/12/18	12/02/18	100%

Workflow

- Choose Prediction Model of Interest
- Create Questions from given medical codes used in model
- Order questions by impact to prediction (reordering if necessary)
- Present questions to user
- Update predicted outcome and confidence level as questions are answered

Architecture



Design

- Read R models (.rds)
- Process Covariate data
- Sort Questions
- Perform Monte Carlo Estimations
- Extendible: Ability to add new PLP models through simple configuration steps

Sorting Covariate Questions

- The top equation determines the weight of evidence
- The bottom equation is for determining the expected weight of evidence on a specific health state
- Questions are sorted by descending expected weight (EW) value

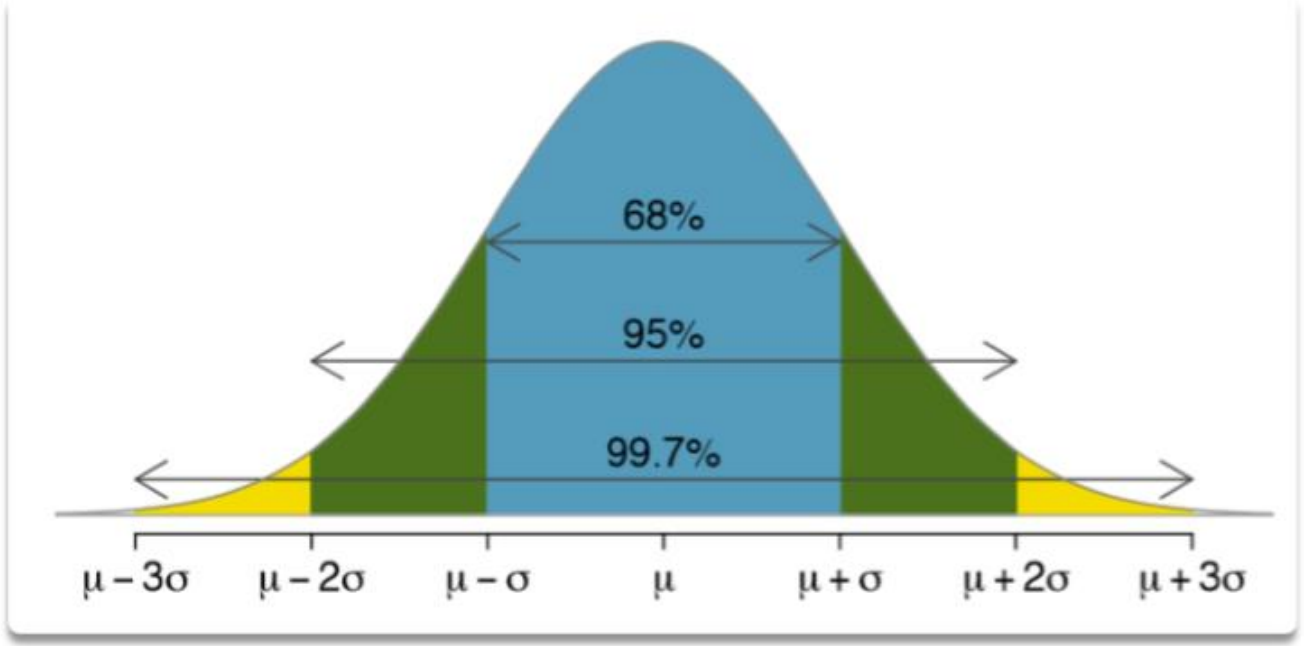
$$W(H:E) = \log\left(\frac{Pr(E|H)}{Pr(E|\bar{H})}\right)$$

$$EW(H:E) = \sum_{k=1}^{m} W(H:t_k)Pr(t_k|H)$$

Monte Carlo Estimation

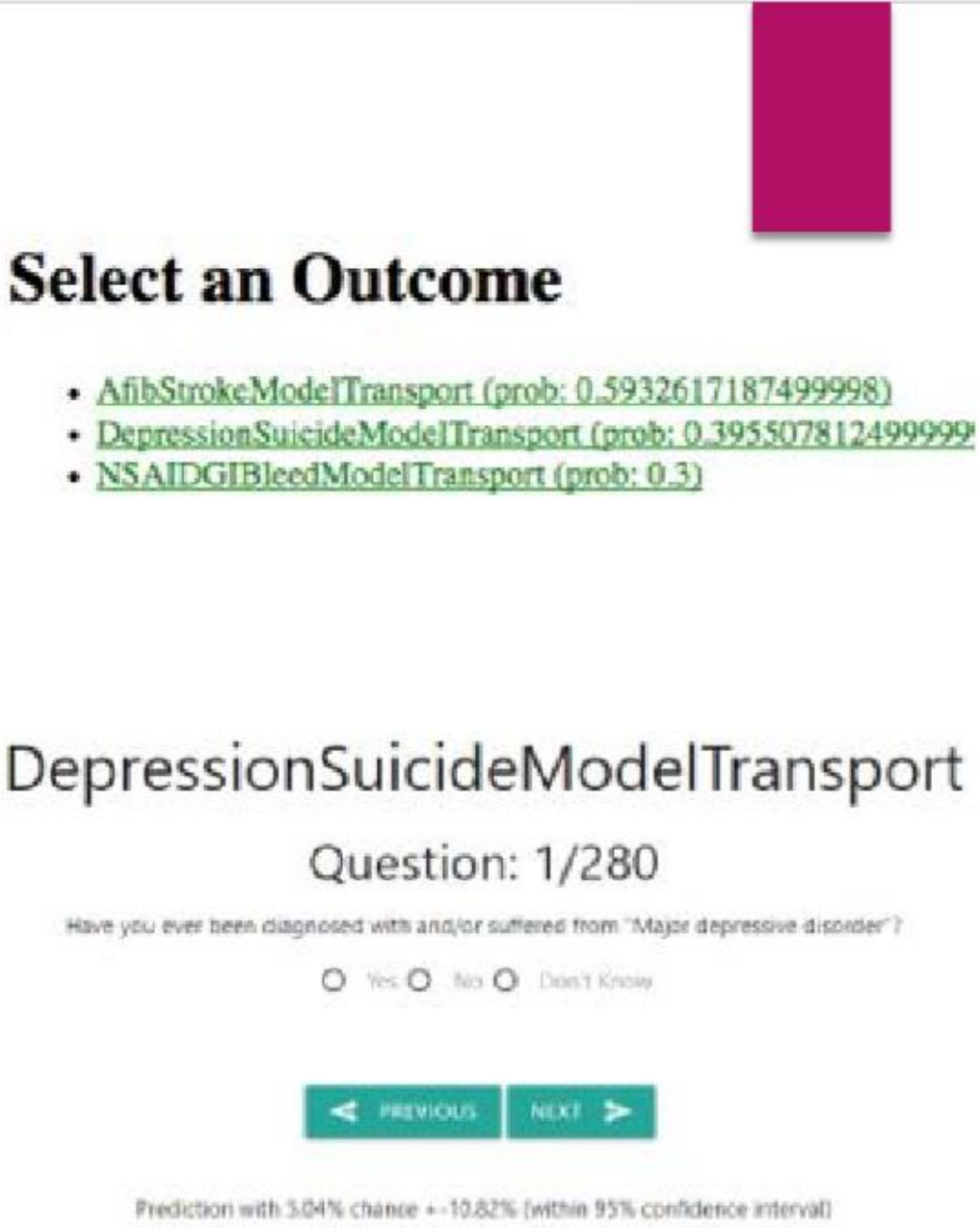
Steps to complete Monte Carlo Estimation:

- Simulate the missing covariate data by covariates marginal probability
- Run logistic regression on simulated covariate outcome from step 1
- Stores outcome and repeats steps 1 and 2 a total of ~2000 times
- Returns mean and standard deviation of stored outcomes. These are used to calculate the 95% interval.



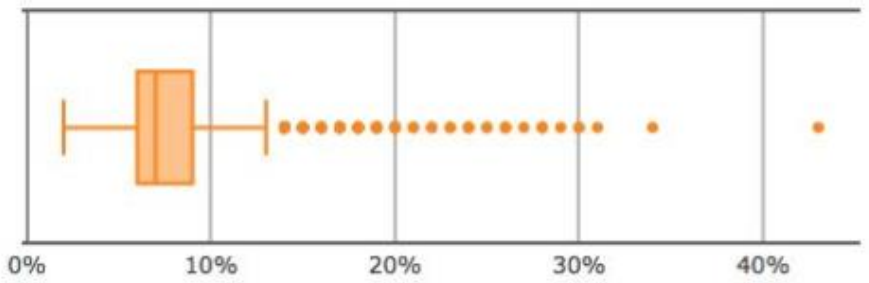
Development Phase

- Developed in Django
- 3 Pre-trained R models from OHDSI
- Rpy2 library
- Backend and Frontend Development

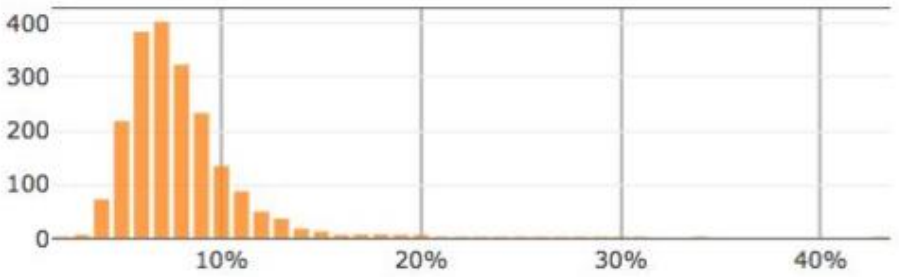


Polish Phase

- Gathering Feedback
- Round 1: too many pages
- Round 2: too many questions
- Added css and js



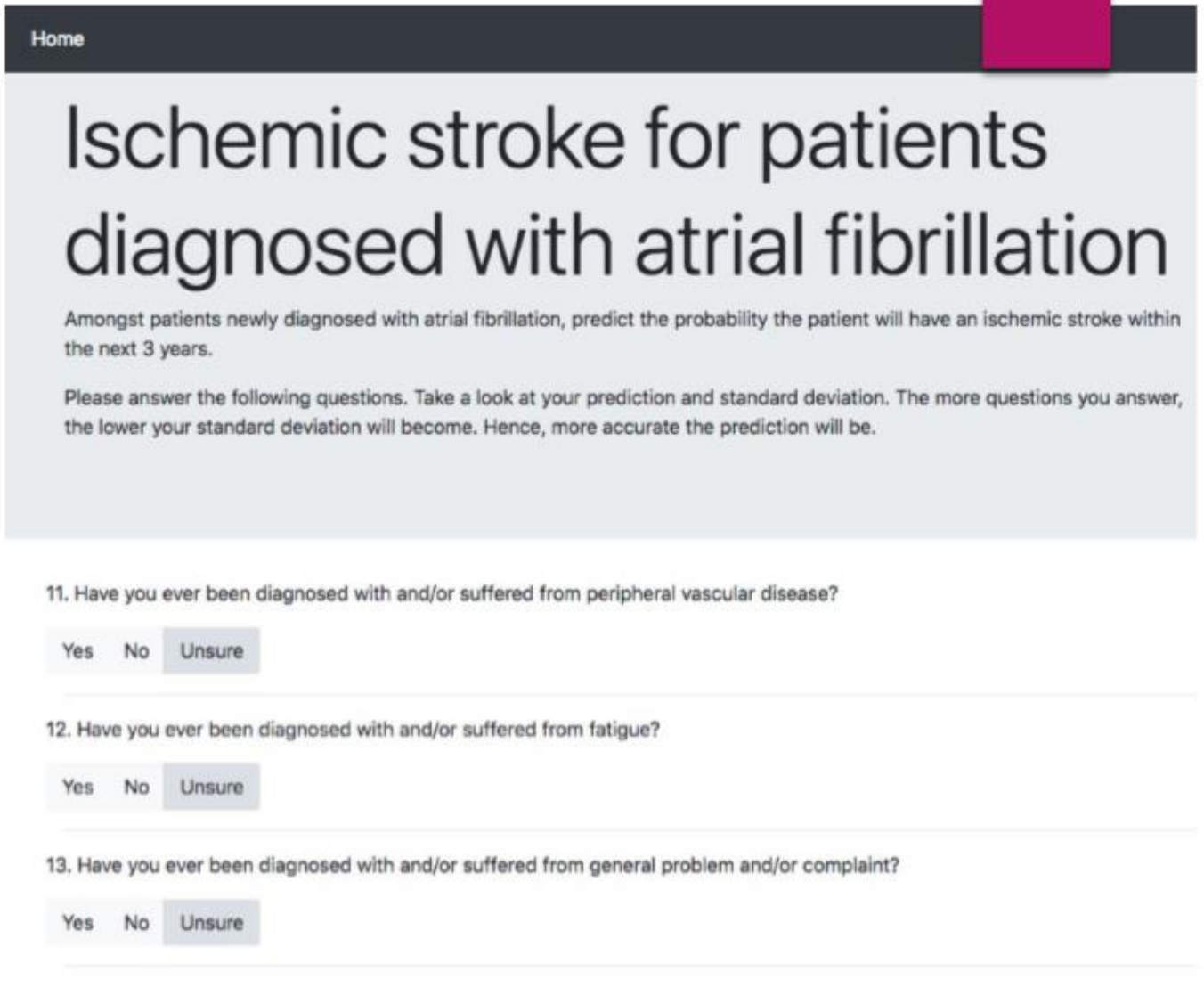
prediction boxplot



prediction histogram

Testing Phase

- Bug fixing
- Storing sessions



DEMO

HTTPS://CS6440-F18-PRJ14.APPS.HDAP.GATECH.EDU/

https://drive.google.com/drive/folders/1wcRsUIAIXXxsUMLRLW_vP-sIXz7bzL8V?usp=sharing

Links/References

- Project Public URL: https://cs6440-f18-prj14.apps.hdap.gatech.edu/
- PLP objects description: https://github.com/OHDSI/PatientLevelPrediction
- 3 example PLP models: https://drive.google.com/drive/folders/1wcRsUIAIXXxsUMLRLW_vP-sIXz7bzL8V?usp=sharing
- Gantt Chart: https://docs.google.com/spreadsheets/d/1FNK3MJTs4e05rsHUXeUyAzfJcM83Is52t-Lmf0Ywhqo/edit#gid=0

Links/References: Continued

- PLP model JAMIA paper: https://academic.oup.com/jamia/article/25/8/969/4989437
- Project Git repository: https://github.gatech.edu/gt-cs6440-hit-fall2018/Consumer-facing-application-for-patient-level-prediction

Thank You!

Questions?