Healthcare AI Suite

Mid Project Review Deck

Naomi Kaduwela
May 15, 2019

Accomplishment Highlights

Top 2-3 things you accomplished this "sprint" (can be analyses, data wrangling, visualizations, interesting things about the data you found out, etc)

- **Data**: Identified data for AI Doctor and AI Cardiologist.
  - Wrote web scraping in python to get symptom disease data from table
  - Found data on kaggle for AI Cardiologist
  - Completed data cleaning for both
- **Model**:
  - First iteration of Decision Tree for Symptom/Diagnosis
  - First iteration of SVM model for Heart Disease classification

- **UI**:
  - Found Avatars for Doctor, Cardiologist, & Eliza
  - Completed interface for symptom selection
  - Completed interface for patient questionnaire too intake required features for heart disease classification model
  - Completed Eliza integration
- **DB**: Created DB model and landed test row of data in RDS & SQLite DB
- **AWS**: EC2 instance set up, S3 bucket set up

Review Progress: Epic 1 - Doctor

What stories from your sprint plan you completed

Theme 1: Create an AI doctor that can suggest the most likely disease based on your list of symptoms.

- Epic 1: Create an analytic that can suggest the most likely diagnosis based on user specified symptoms, along with a confidence interval
    - US0: Environment setup: install any required packages, libraries. Spin up EC2 on AWS. Create requirements.txt to make reproducible environment.
    - US1: Find a good data set for training, including: list of symptoms, disease, and the count of disease occurrence.
    - US2: Complete data transformations in python so the data can be fed into the model
    - US3: Create a AWS RDS to store user input from UI and model output
    - US4: Build decision tree based classification model in python

Review Progress: Epic 2 - Cardiologist

What stories from your sprint plan you completed

Theme 2: Create an AI doctor that can diagnose heart disease based on patient health factors

- Epic 2: Create an analytic that can suggest the most likely to diagnose heart disease based on patient health factors
  - US0: Environment setup: install any required packages, libraries. Spin up EC2 on AWS. Create requirements.txt to make reproducible environment.
  - US1: Find a good data set for training, including: list of features of heart disease
  - US4: Build SVM classification model in python
 - US5: Build UI of cardiologist doctor

Review Progress: Epic 3 - Eliza

Theme 2: Create an AI therapist. It has been found to be calming and this would be helpful as users might be distressed as they are receiving their diagnoses.

- Epic 2: Integrate with ELIZA therapy chat-bot with sentiment detection and response via avatars so that there is a comforting AI to listen to users, without judgement.
    - US1: Integrate ELIZA Javascript codes: Elizabot.js, elizadata.js, setimood.js
    - US2: Integrate ELIZA images for avatar: happy, sad, excited, etc
    - US3: Create custom HTML/CSS file to visualize Eliza therapy bot chat next to the symptom diagnoser
    - US4: UAT test scripts to test ELIZA is properly integrated into the UI and all functionality is working as expected
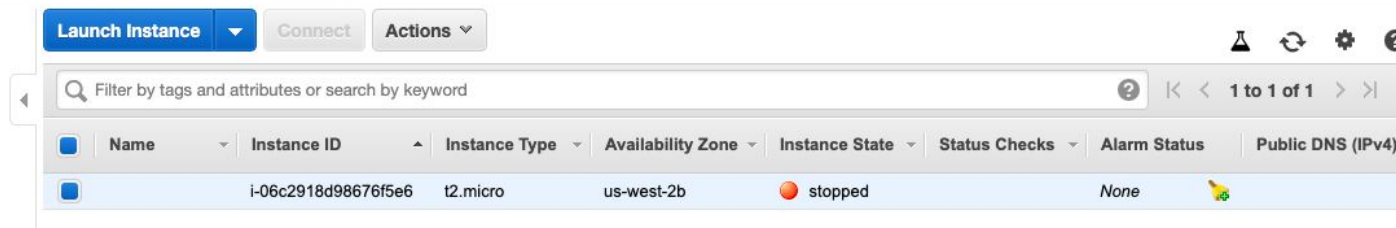
AWS Demo

S3 Bucket:

EC2:

RDS:

Model Demo

AI Cardiologist - SVM:

AI Doctor - Decision Tree:

Testing
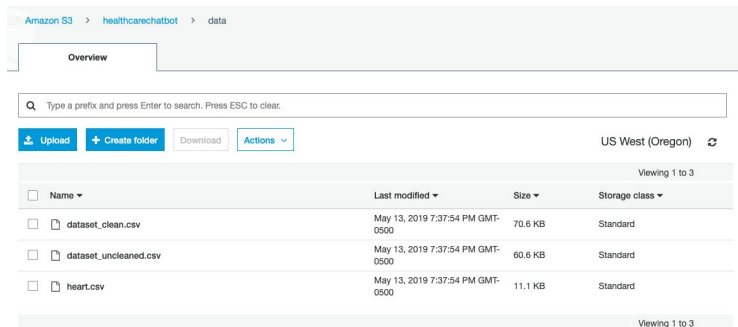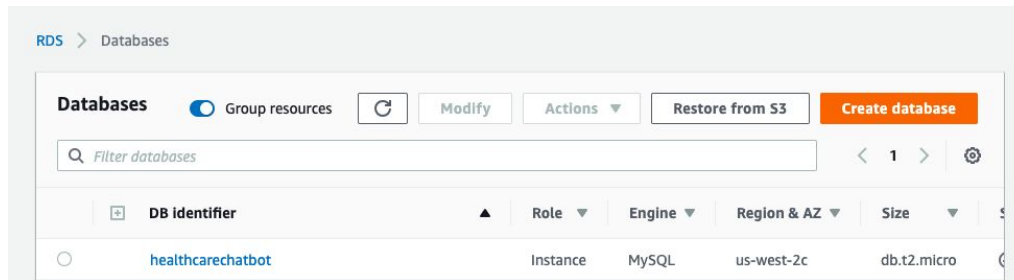
```
1  from sklearn.metrics import classification_report, confusion_matrix
2  print(confusion_matrix(y_test,y_pred))
3  print(classification_report(y_test,y_pred))
```

```
[[39  3]
 [ 4 54]]
              precision    recall  f1-score   support

           0       0.91      0.93      0.92        42
           1       0.95      0.93      0.94        58

   micro avg       0.93      0.93      0.93       100
   macro avg       0.93      0.93      0.93       100
weighted avg       0.93      0.93      0.93       100
```
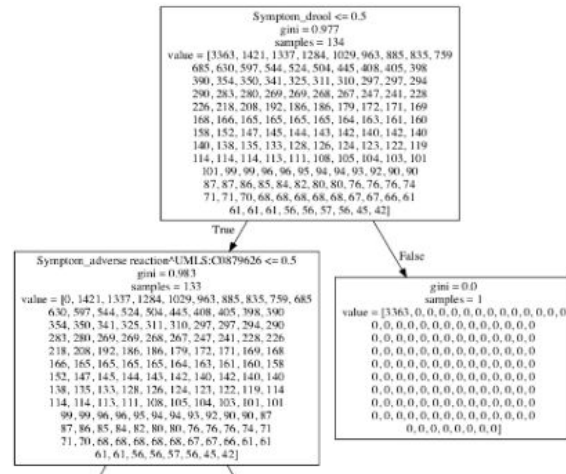
UI Demo - Landing Page

Good Morning, Naomi!

Which doctor would you like to see today..?

| | Home | | General Doctor | | Cardiologist | | Psychologist | | Neurologist | |

Dr. Bot

Dr. Bot is a **Decision Tree algorithm** that will output the most likely disease based on a patient's symptoms.

Dr. Heart

Dr. Heart is a **Support Vector Machine algorithm** that can predict the likelihood of heart disease based on 14 features.

Dr. Eliza

Dr. Eliza is a **Chat Bot** that can converse with patients to help calm their anxieties and make them feel at ease.

Dr. Brain

Dr. Brain is a **Convolutional Neural Network** that will classify MRI scans as positive or negative for brain tumors.

UI Demo

Dr. Heart

Patient Admission

Full Name *	First Last
	Invalid value
Age *	
Gender *	Female
Chest Pain Type *	2 = atypical angina
	displays the type of chest-pain experienced
Resting Blood Pressure *	
	mmHg (unit)
Serum Cholestrol *	
	in mg/dl (unit)
Resting Blood Pressure *	
	mmHg (unit)
Fasting Blood Sugar *	2- blood sugar < 120mg
	compares the fasting blood sugar value of an individual with 120mg/dlc
Resting ECG *	1 = having ST-T wave ab
	compares the fasting blood sugar value of an individual with 120mg/dlc
Max Heart Rate *	
Exercise induced angina *	-Select-
ST depression induced by exercise relative to rest *	
	Integer or Float value
Peak exercise ST segment *	-Select-
Number of major vessels (0-3) colored by flourosopy *	
	Integer or float
Thal *	
	displays the thalassemia

Submit

Dr. Bot

Please select your symptoms:

☐ Headache   ☐ Nausea   ☐ Bleeding   ☐ Fainting   ☐ xyz

Please select your symptoms:

pain chest
shortness of breath
dizziness
asthenia
fall
syncope
vertigo
sweat
sweating increased
palpitation
nausea
angina pectoris
pressure chest
polyuria
polydypsia
orthopnea
rale
unresponsiveness
mental status changes
vomiting
labored breathing
feeling suicidal
suicidal

Selected Symptoms:

View Decision Tree

Eliza+

ELIZA: Eliza here. Is something troubling you?
YOU: Hi Eliza, I just found out I might have cancer.
ELIZA: That's quite interesting.
YOU: Is it?
ELIZA: Tell me more about that.
YOU: The doctor told me and I was shocked.
ELIZA: Why do you tell me you were shocked now?
YOU: Because I had no idea, there were no symptoms.
ELIZA: You say because you had no idea?
YOU: That's right
ELIZA: Does talking about this bother you?
YOU: a little bit
ELIZA: Tell me more about that.
YOU: acknowledging it makes it true
ELIZA: That is interesting. Please continue.

Talk   Reset

Lessons Learned

1. **Data**: Cannot make predictions at an individual level from aggregate data. Thus, the underlying data you feed into the model is key and needs to be ensured it will answer the specific question you want from your data.
2. **Model Training**: Testing in decision tree is complicated to do CV when you need to feed all the data into the model to train it you cannot find a way to test it unless you generate data, but this gets messy and has its own limitations.
3. **Overall**: Since there were many learning with the doctor decision tree model, though it will not be removed (as the UI looks good and it functions as a prototype), will focus more on the new cardiologist model.

Recommendations & Next Steps

1. Break apart python notebooks into ingestion, EDA, model development, testing, etc
2. Write unit test scripts
3. Create scripts to select data from UI and feed it to the model and then return the output to the UI
4. Write UAT test scripts
5. Create FLASK app

6. If time permits, integrate the Convolutional Neural Net we built to classify brain tumors in a Dr. Brain page!

Thank you!