



CHECK FIRST

(Analyze COPD and Dementia with us.....)

Project manager : Andy Vu
Team: Harshita Panchal
Emily Blackstein
Amar Sen

Date : September 07, 2023

Overview

Focus of this project is to use Machine Learning and Deep Learning models to determine the risk of common health diseases.



COPD

DIMENTIA

(Chronic Obstructive Pulmonary Disease)

Overview:

"Check First" is an app created specifically to analyse the health-related risk behaviors, chronic health conditions, and use of preventive services. We are analyzing risk for COPD(Chronic Obstructive Pulmonary Disease) and Dementia.





COPD (Chronic Obstructive Pulmonary Disease)

Chronic obstructive pulmonary disease (COPD) is a lung disease and makes it hard to breathe.

*** Risk factors:**

- **Smoking**
 - **Exposure to certain gases or fumes in the workplace**
 - **Frequent use of a cooking fire without proper ventilation**
 - **Exposure to heavy amounts of secondhand smoke and pollution**

*** Symptoms :**

- **Cough, with or without mucus**
 - **Fatigue**
 - **Frequent respiratory infections**
 - **Shortness of breath (dyspnea) that gets worse with mild activity**
 - **Trouble catching one's breath**
 - **Wheezing**



Dementia

Dementia is a general term for loss of memory, language, problem-solving and other thinking abilities that are severe enough to interfere with daily life.

*Causes

- Depression.
- Medication side effects.
- Excess use of alcohol.
- Thyroid problems.
- Vitamin deficiencies.

*Symptoms

- Short-term memory.
- Keeping track of a purse or wallet.
- Paying bills.
- Planning and preparing meals.
- Remembering appointments.
- Traveling out of the neighborhood.



Following Models were trained and compared for COPD

(PREDICTIVE MODEL)

1st Model:

Support Vector Machine algorithm, C-Support Vector Classification.

2nd model:

k-nearest neighbors algorithm, k-nearest neighbors vote.

3rd model:

Ensemble-based methods for classification, Gradient Boosting Classifier.

4th model:

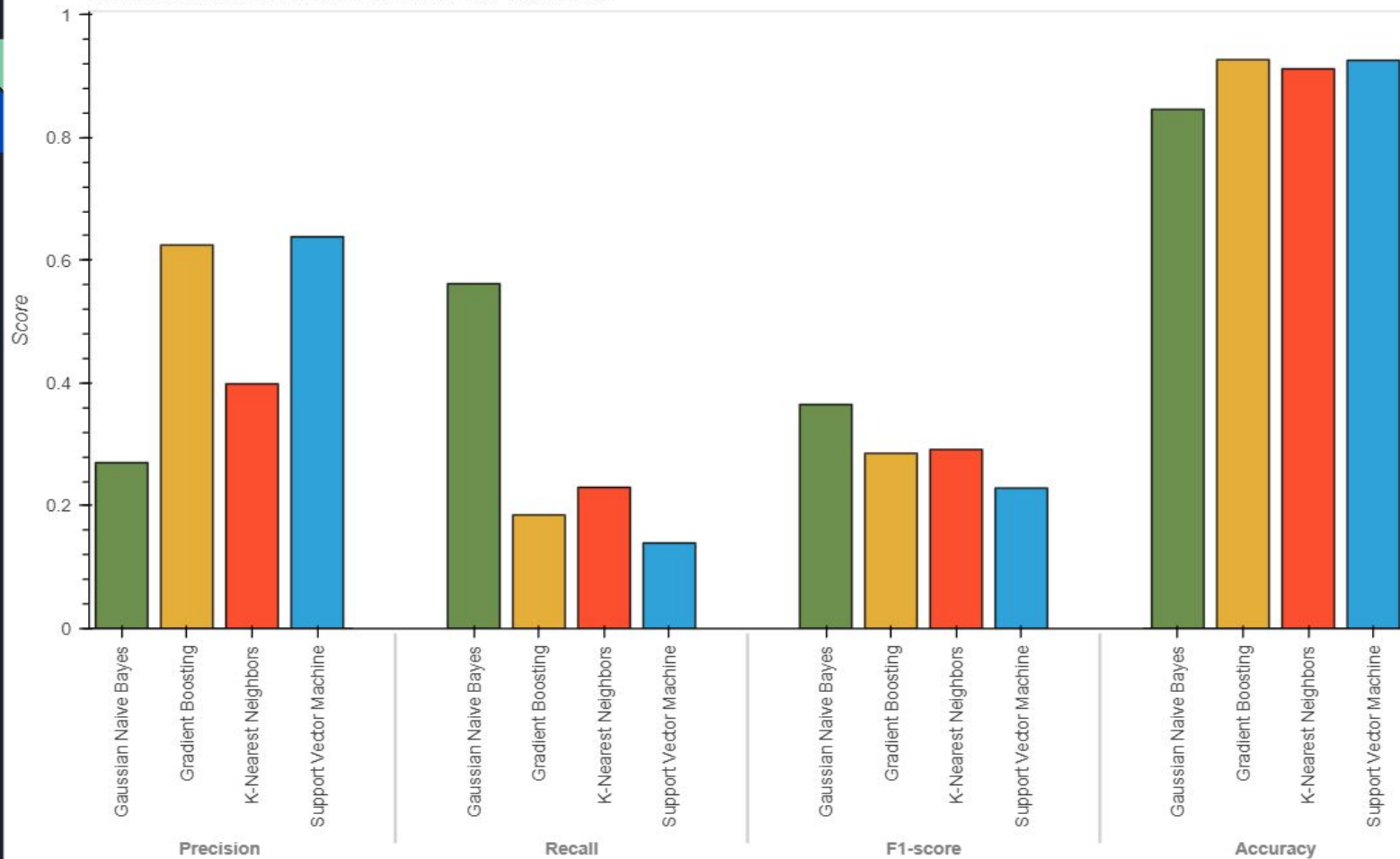
Naive Bayes algorithm, Gaussian Naive Bayes Classifier.



Data Presentation & Model Training Process

1. Clean the Data
2. Seperate the label and features
3. Split the data into Training and Testing sets
4. Scaled the data to standardize
5. Fit the Model
6. Make Predictions
7. Performance Evaluation

ML Model Scores for Positive for C.O.P.D.





Following Models were trained and compared for Dementia

1st Model:

Support Vector Machine algorithm, C-Support Vector Classification.

2nd model:

k-nearest neighbors algorithm, k-nearest neighbors vote.

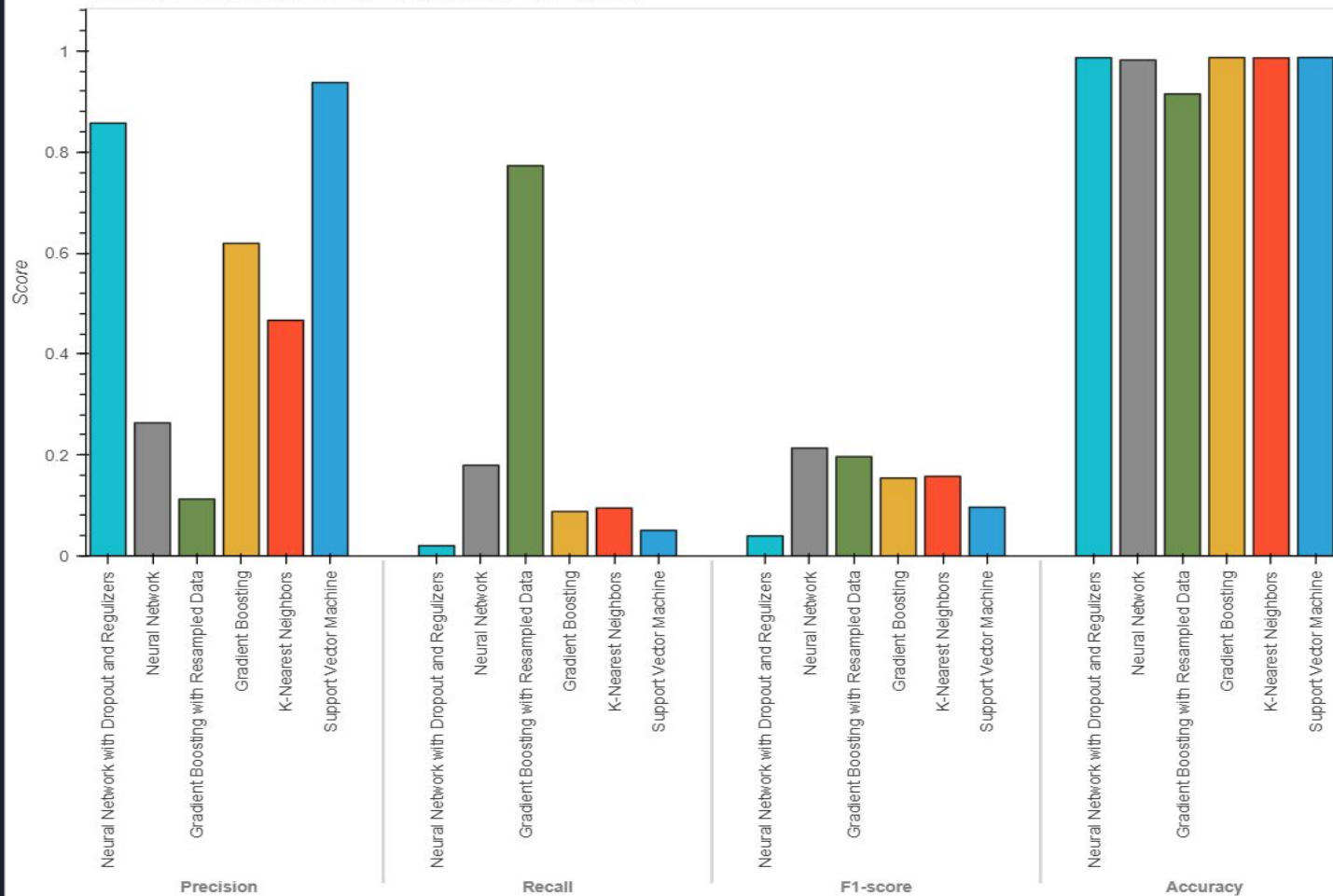
3rd model:

Ensemble-based methods for classification, Gradient Boosting Classifier.

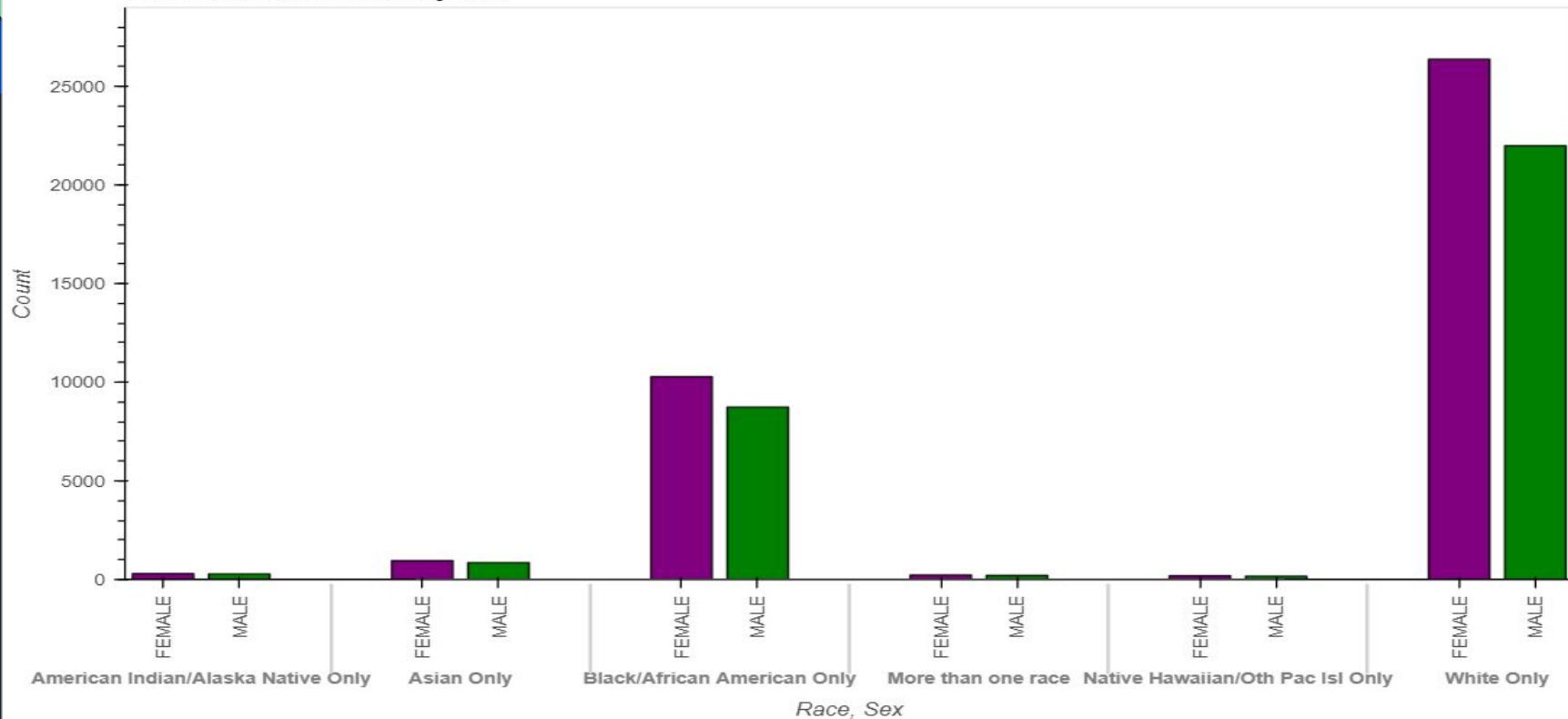
4th Model:

Neural Network

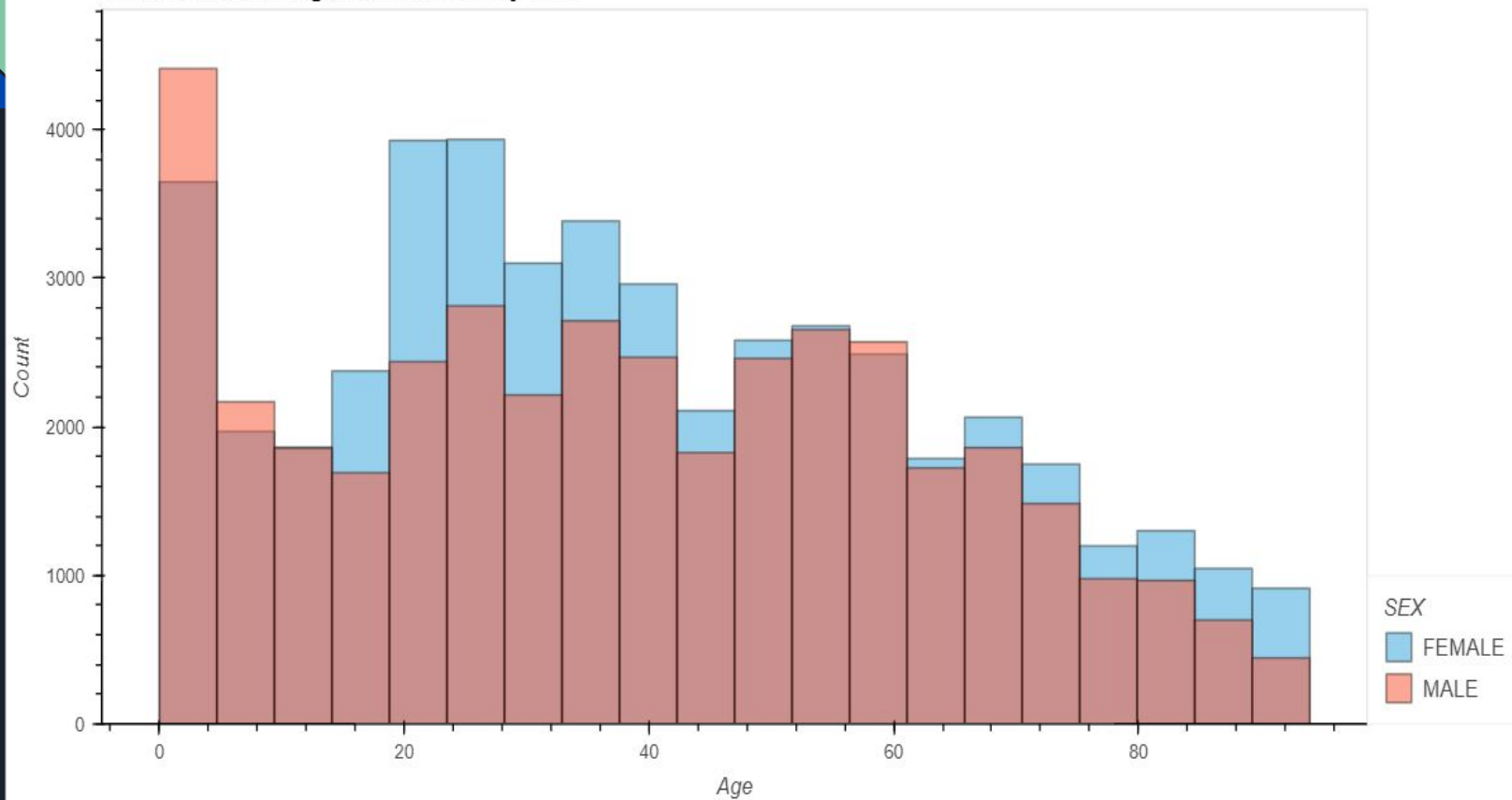
ML/NN Model Scores for Positive for Dementia



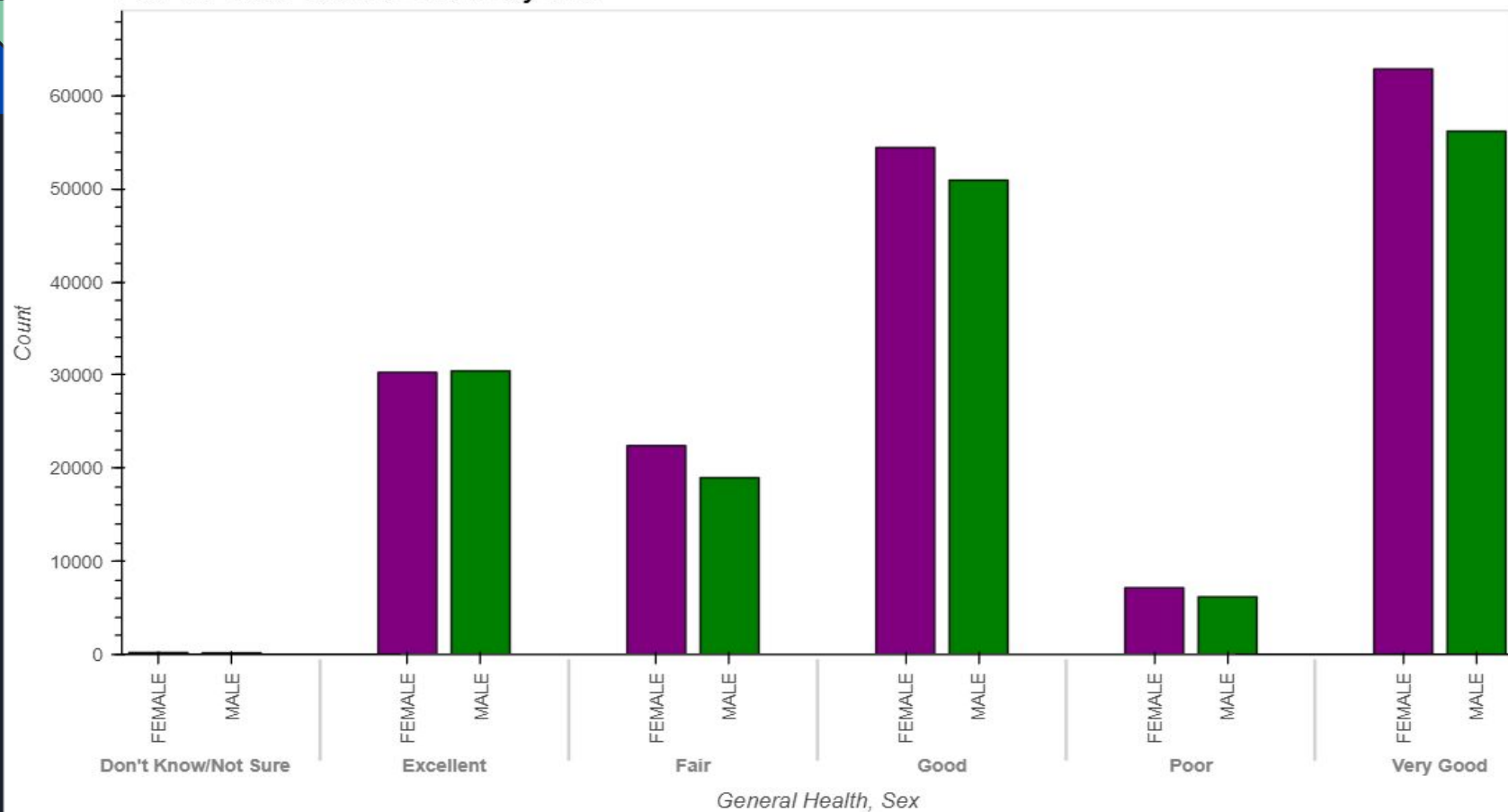
NHAMCS Data: Race by Sex



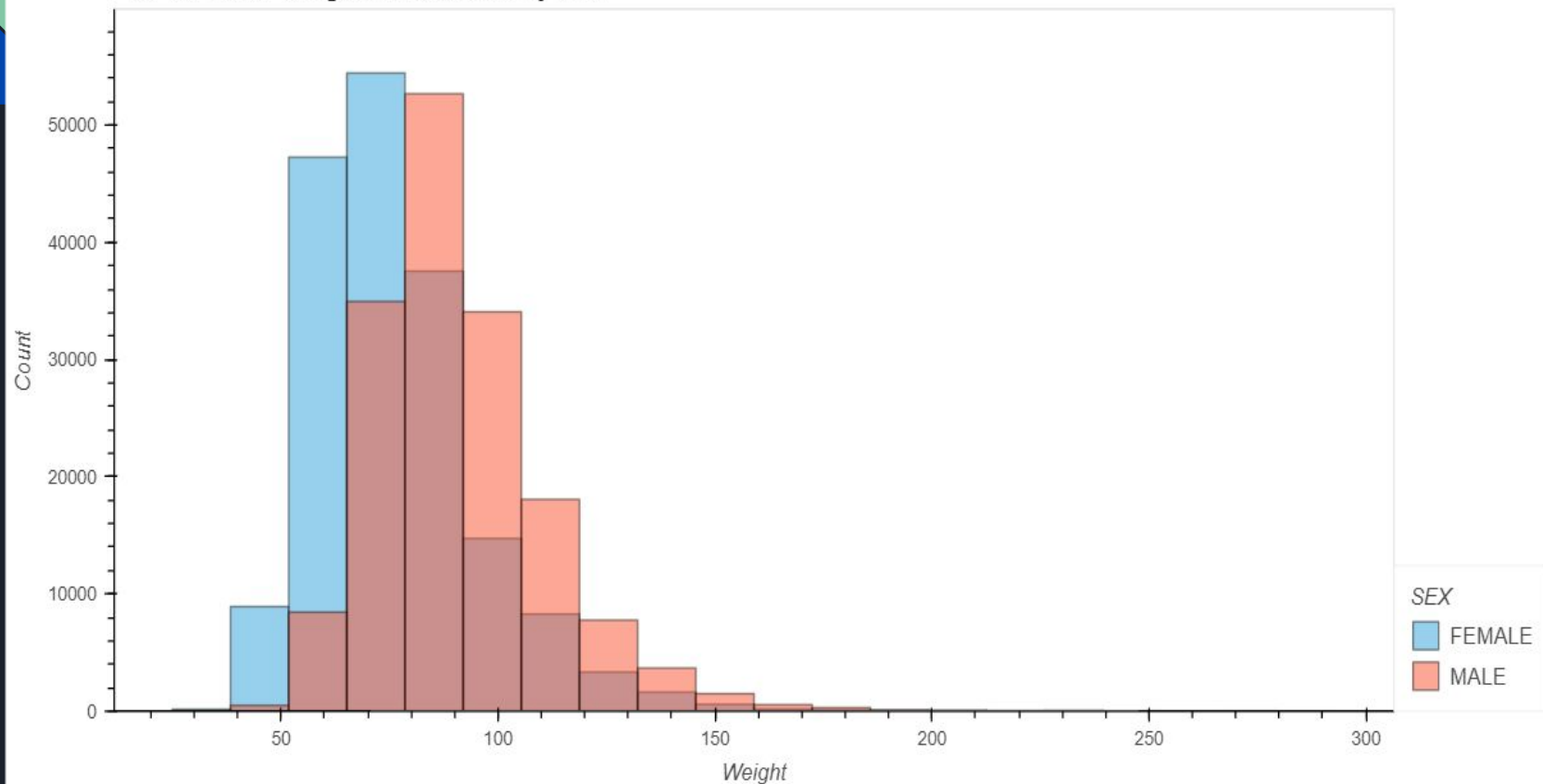
NHAMCS Data: Age Distribution by Sex



BRFSS Data: General Health by Sex



BRFSS Data: Weight Distribution by Sex



Conclusion & Predictions - COPD

	Support Vector Machine	K-Nearest Neighbors	Gradient Boosting	Gaussian Naive Bayes
Precision	0.638414	0.398810	0.625000	0.270344
Recall	0.139445	0.230069	0.185130	0.561959
F1-score	0.228894	0.291801	0.285648	0.365065
Accuracy	0.926011	0.912055	0.927081	0.846061

Conclusion & Predictions - Dementia

	Support Vector Machine	K-Nearest Neighbors	Gradient Boosting	Gradient Boosting with Resampled Data	Neural Network	Neural Network with Dropout and Regularizers
Precision	0.937500	0.466667	0.619048	0.112704	0.263682	0.857143
Recall	0.050847	0.094915	0.088136	0.772881	0.179661	0.020339
F1-score	0.096463	0.157746	0.154303	0.196721	0.213710	0.039735
Accuracy	0.987161	0.986339	0.986979	0.914927	0.982181	0.986750

Q & A



Thank
you!

