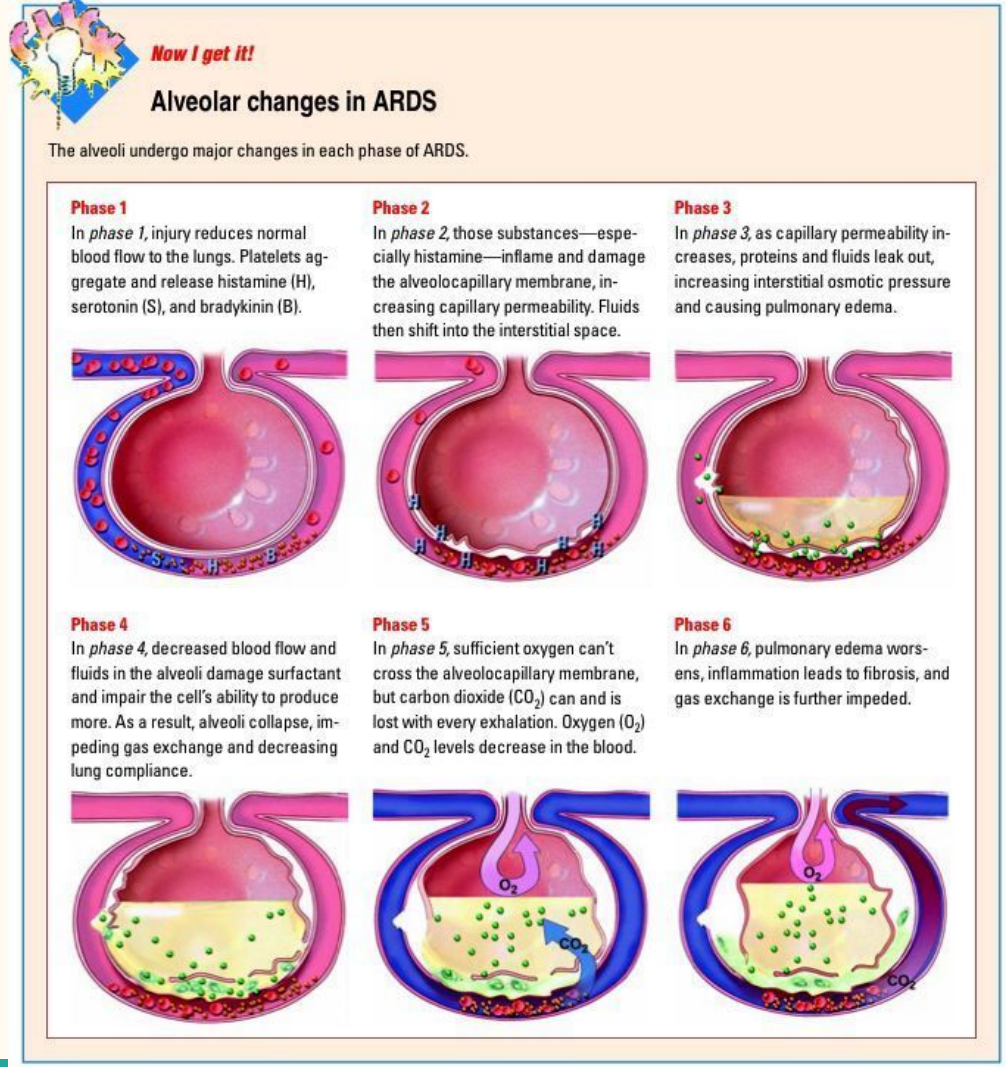
The ARDS Prediction Via Machine Learning

—

Paul, Brihat, Neha, & Saharsh

Introduction

- Inflamed lungs = can't breath
- Drowning, can't breath
- Pre-symptoms; pre-alignments
- 20 - 50 % mortality



Methods

Data Process

- The data was given in encoded using cTAKES
- The data was tokenized, and kept in dictionary using countvectorizer
- Used tfidfTransformer, to standardize the data
- K-fold
- Grid Search

Algorithm

- Multinomial Naive bayes
- LinearSVC
- SVC
- SGD Classification
- Decision Tree Classifier

Results

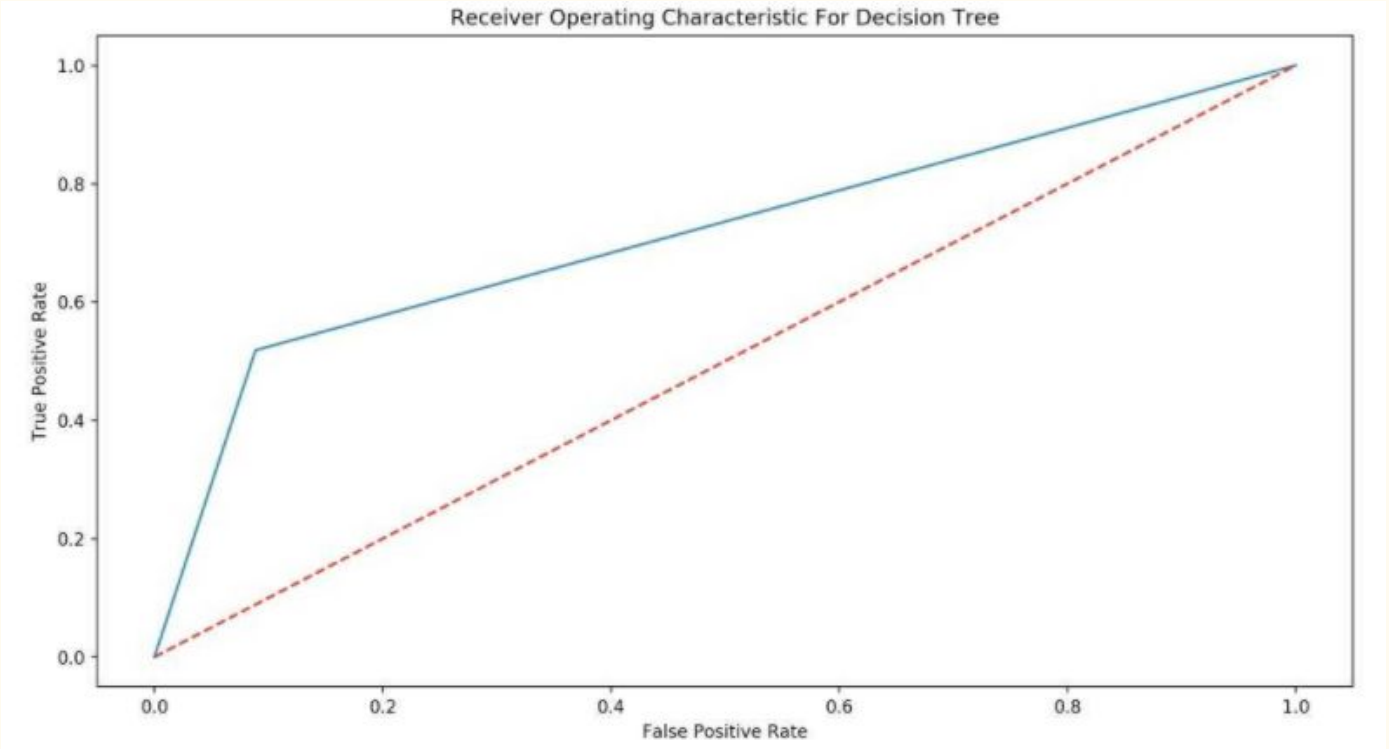
SVC

```
{'clf__C': 10.0, 'clf__gamma': 0.1, 'clf__kernel': 'rbf'}
Accuracy: 0.774
             precision    recall  f1-score   support

          0       0.82      0.89      0.85        79
          1       0.57      0.44      0.50        27

avg / total       0.76      0.77      0.76       106

[[70  9]
 [15 12]]
```

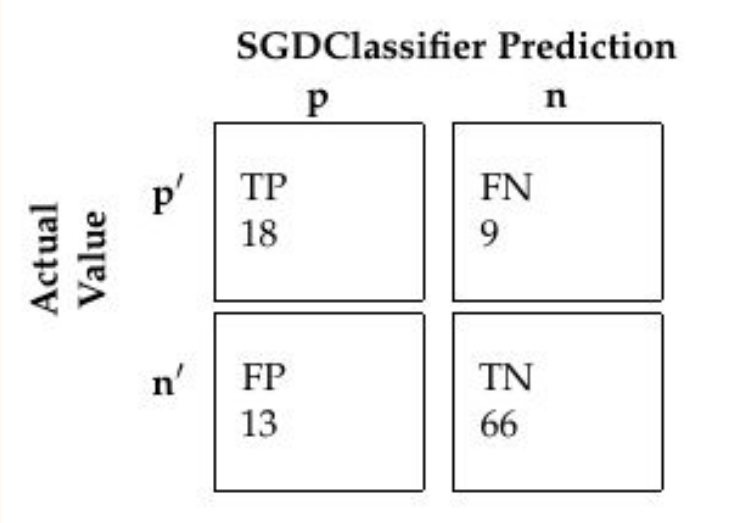
Tree

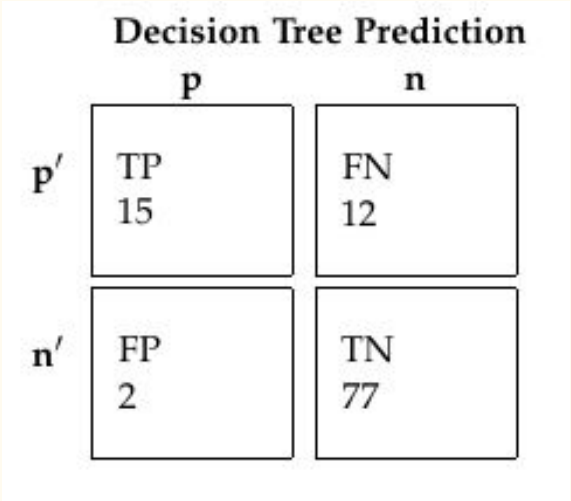
```
{'clf__criterion': 'entropy', 'clf__max_depth': 5, 'clf__max_leaf_nodes': 10, 'clf__min_samples_leaf': 10, 'clf__min_samples_split': 2}
Accuracy: 0.868
             precision    recall  f1-score   support

          0       0.87      0.97      0.92        79
          1       0.88      0.56      0.68        27

avg / total       0.87      0.87      0.86       106

[[77  2]
 [12 15]]
```

ROC Curve for Decision Tree



Confusion Matrix





Conclusion

+ Decision Tree did a better job on average F1 Score
+ F1 score for positive class is better using SGD classifier