External Evaluation of BTC vs Scikit-Learn on Real Data

By: Alex Makhratchev

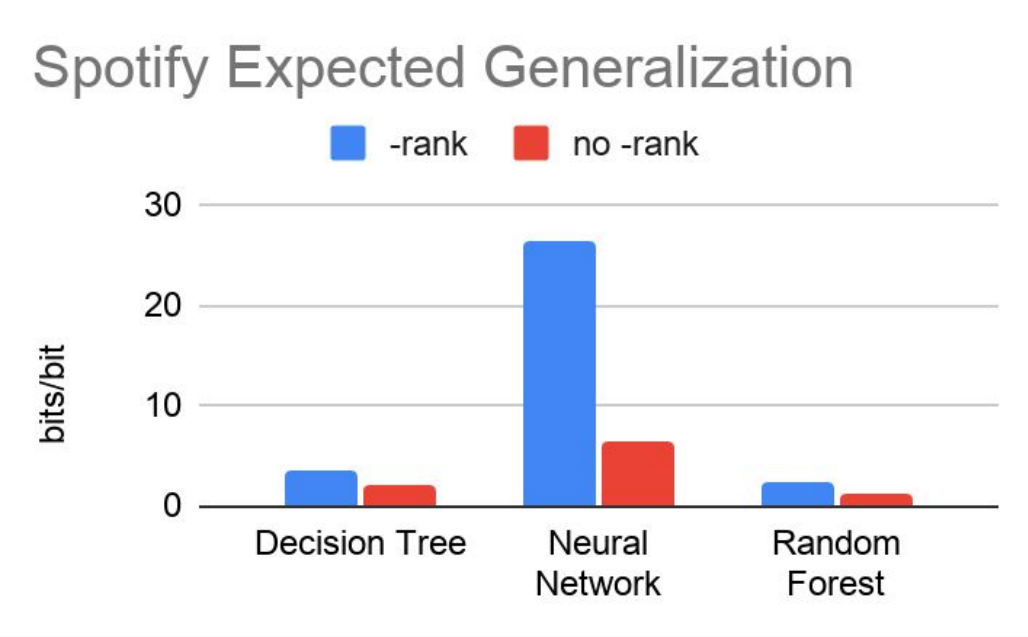
Approach

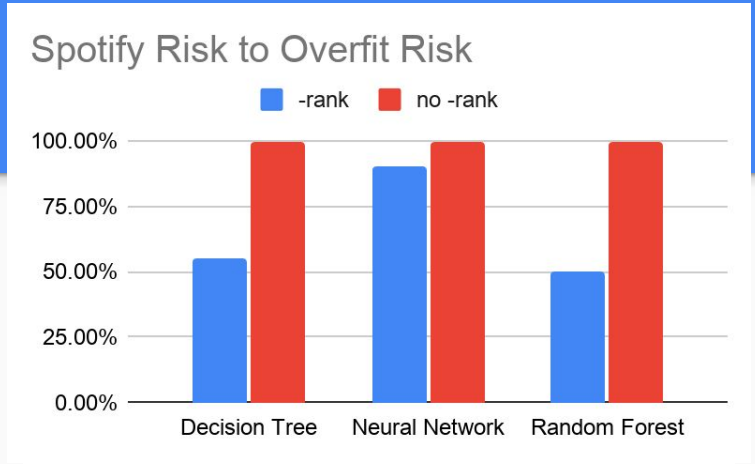
- Scikit Baseline
- Using BTC insights in improve baseline models
- Build Brainome model

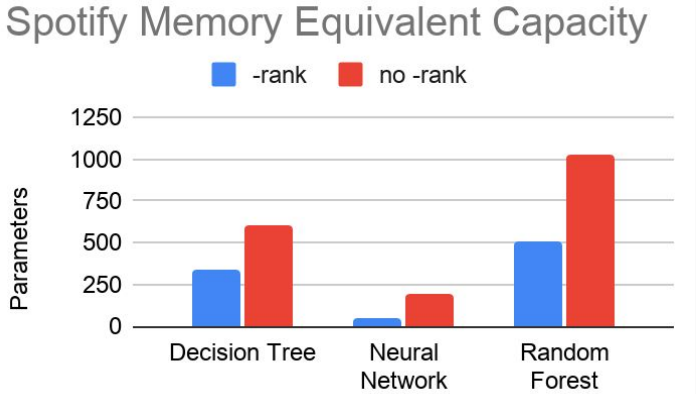
Spotify Music Preference

- Goal: Predict if I will dislike or like a song
- Example attributes: loudness, danceability
- Number of instances: 1224
- Number of attributes: 17
- Number of classes: 2
- Class balance: 43.38%, 56.62%

Spotify Music Preference Measurements

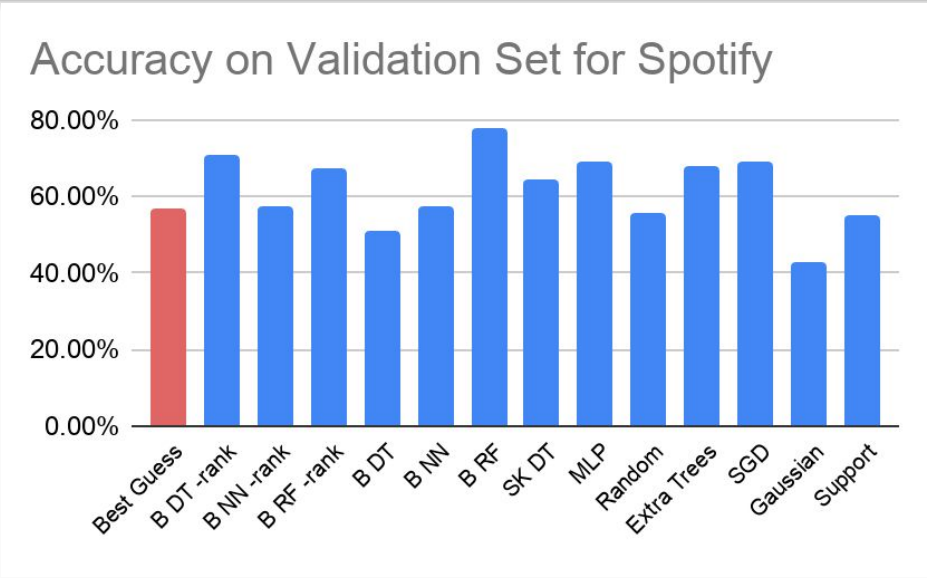


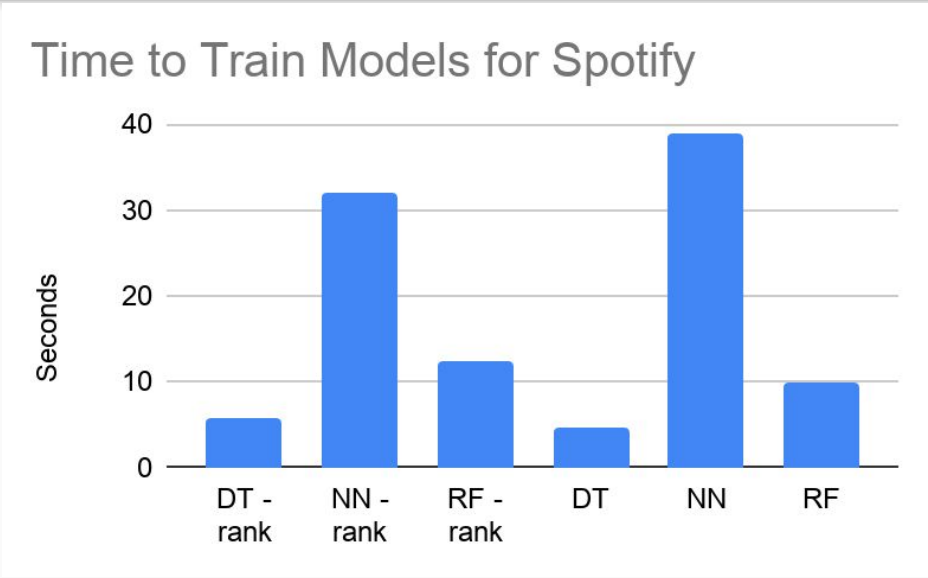




*RF measurements might not be accurate, using alpha of BTC

Spotify Music Preference Model Results





Spotify Summary

Best Model:

- Brainome RF without rank

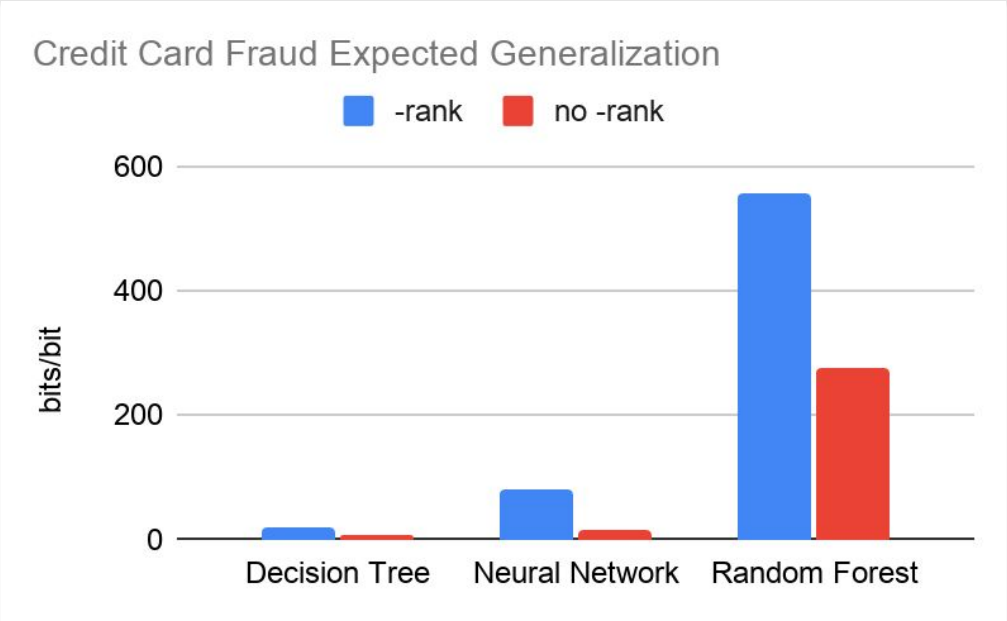
Useful measurement:

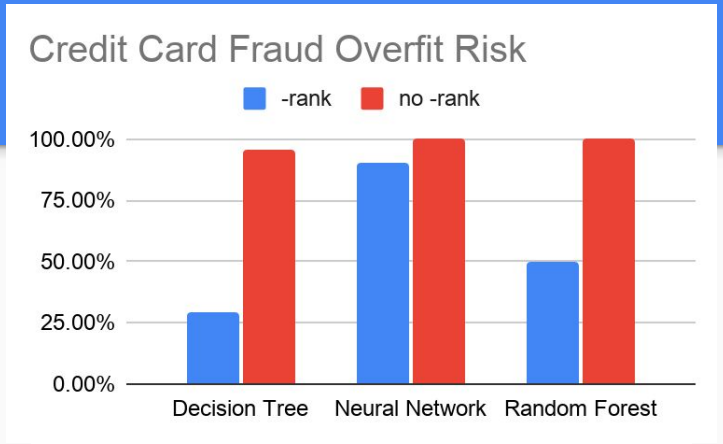
- Using alpha version of BTC, so measurement for RF are inaccurate

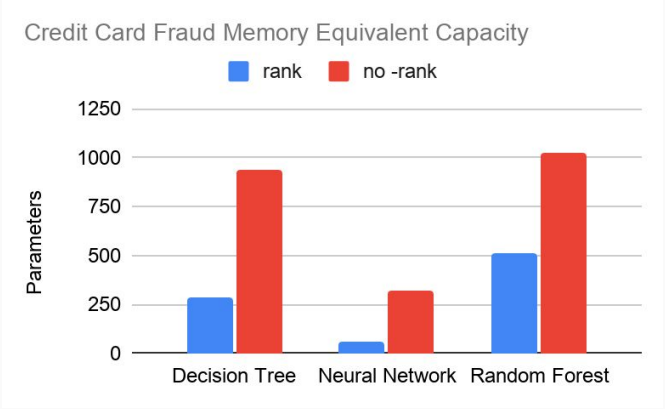
Credit Card Fraud Detection

- Goal: Predict if a transaction is fraudulent
- Example features: time, amount, all other are PCA transformed
- Number of instances: 284807
- Number of attributes: 30
- Number of classes: 2
- Class balance: 99.83%, 0.17%

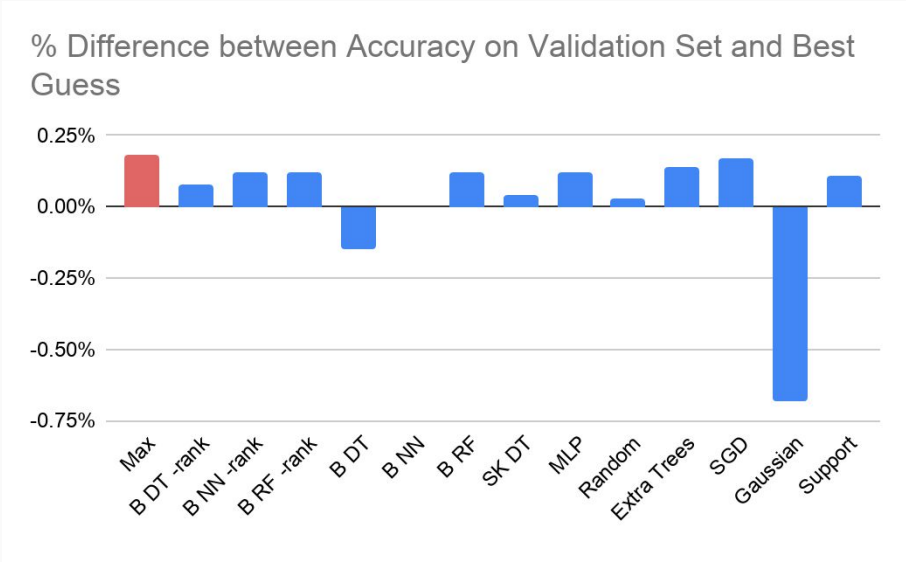
Credit Card Fraud Detection Measurements

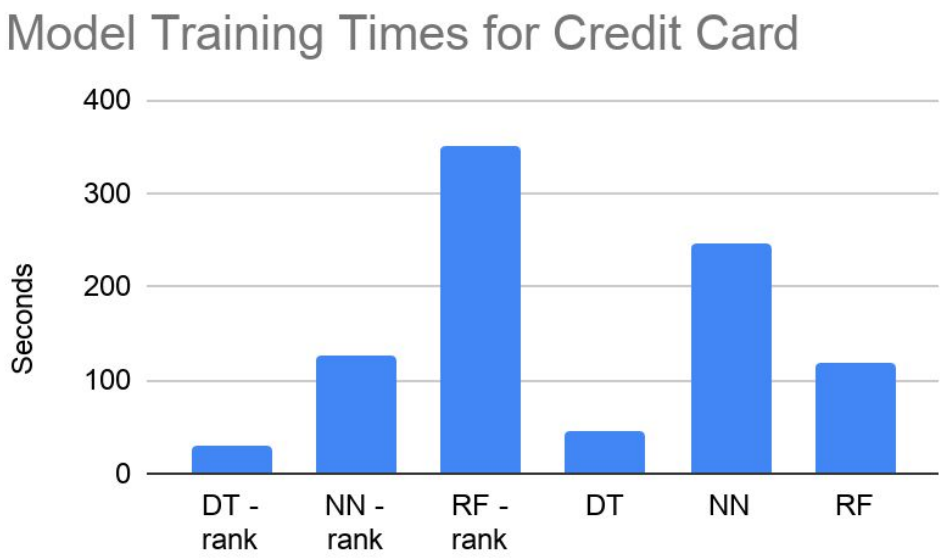






Credit Card Fraud Detection Model Results





Credit Card Fraud Detection Summary

Best Model:

- SK SGD
- Brainome RF -rank and no rank

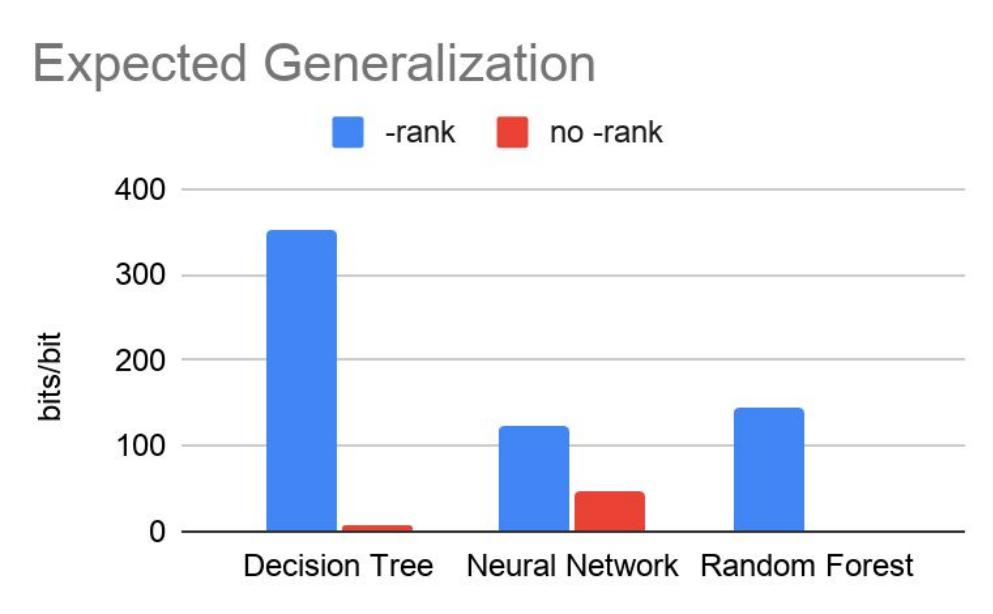
Useful measurement:

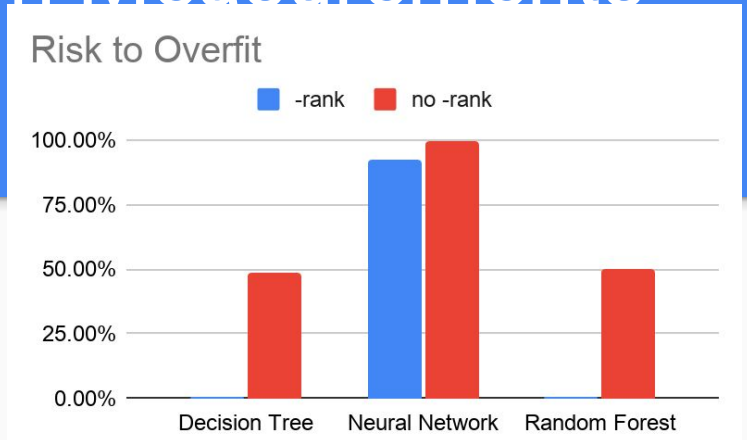
- Highest Generalization

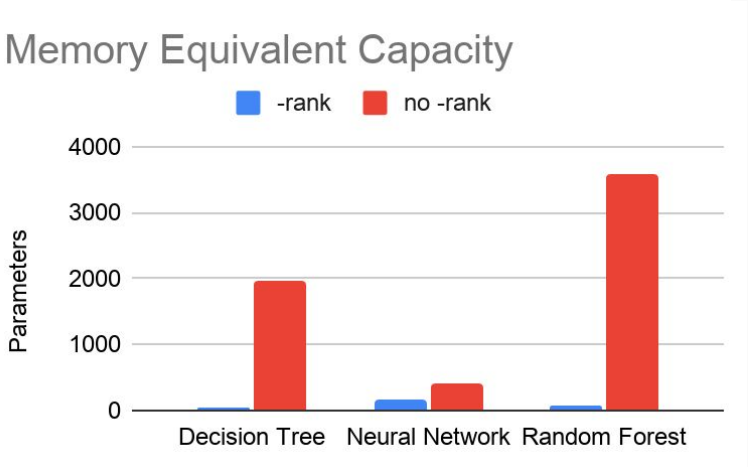
Mushroom Classification

- Goal: classify the type of mushroom
- Attributes: cap-shape, odor
- Number of instances: 8124
- Number of attributes: 22
- Number of classes: 7
- Class balance: 4.53% 26.44% 3.59% 38.75% 14.08% 2.36% 10.24%

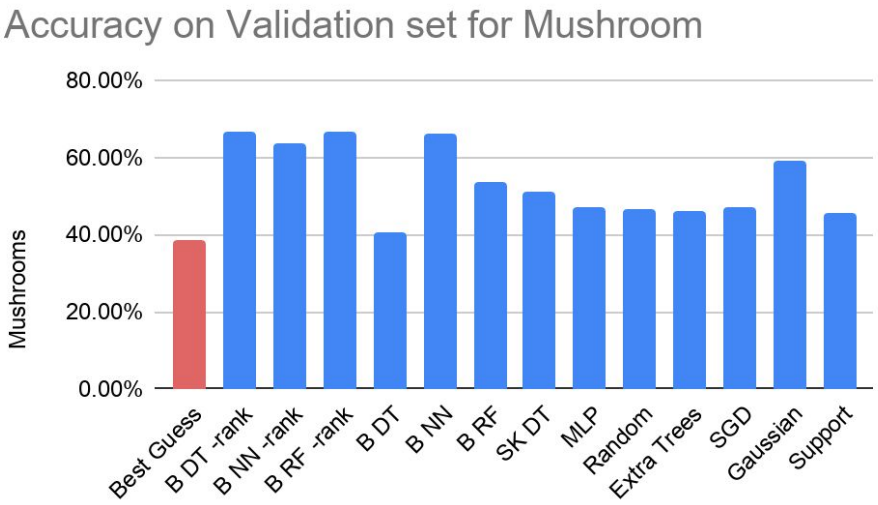
Mushrooms Classification Measurements

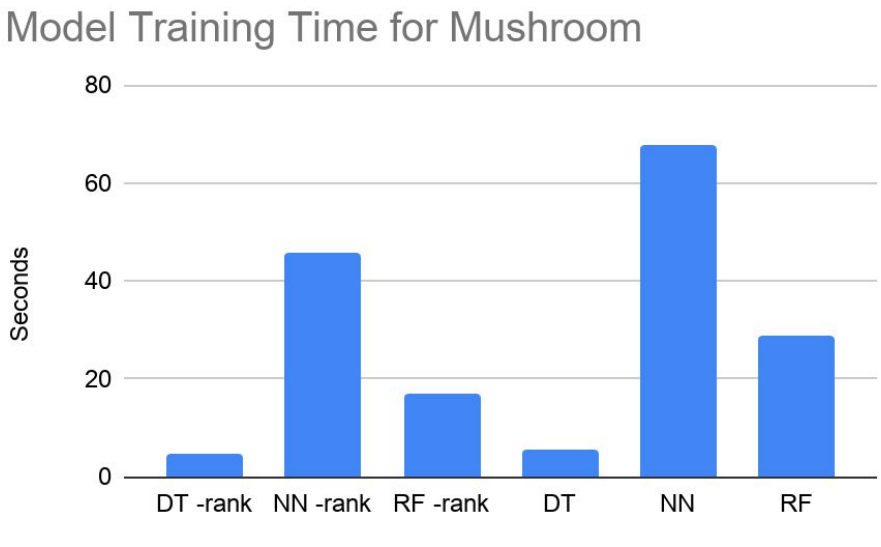






Mushrooms Classification Model Results





Mushroom Classification Summary

Best Model:

- Random Forest -rank and DT -rank

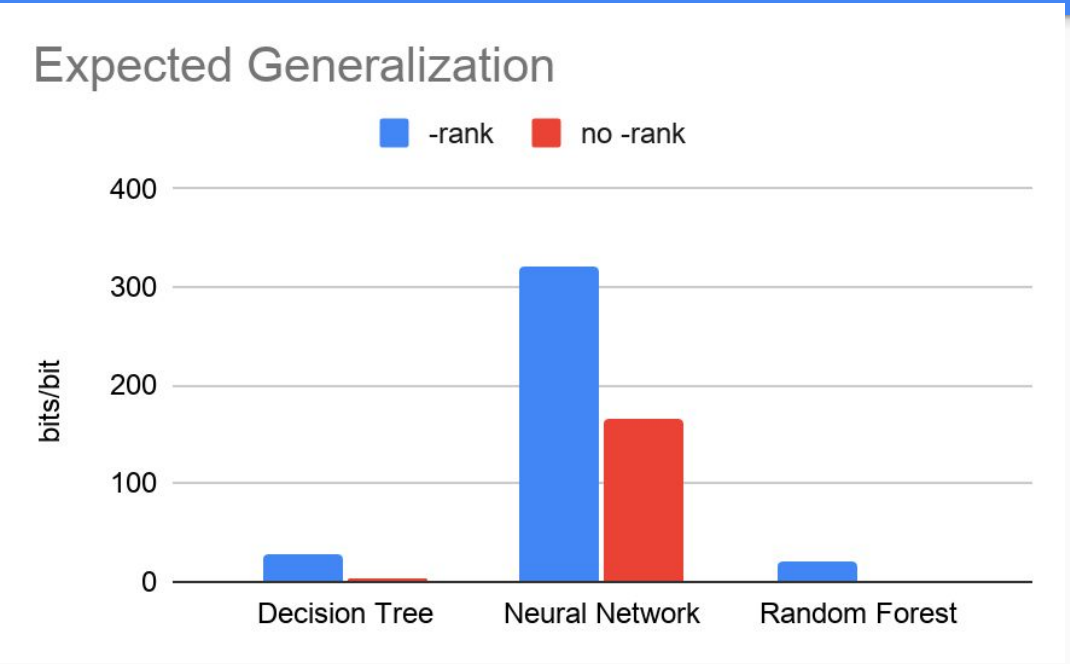
Useful measurement:

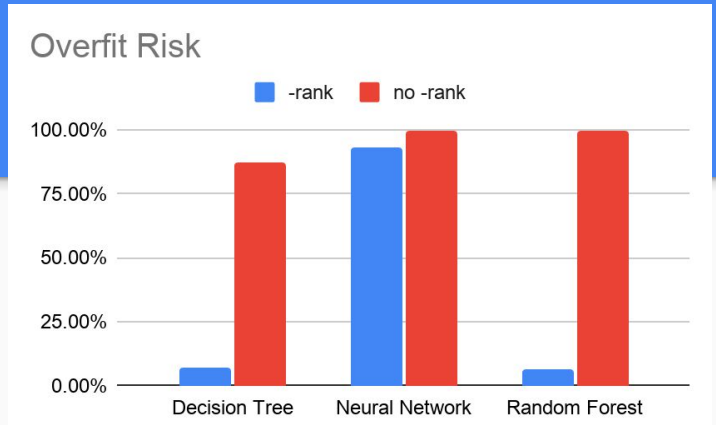
- RF and DT -rank have Highest Generalization

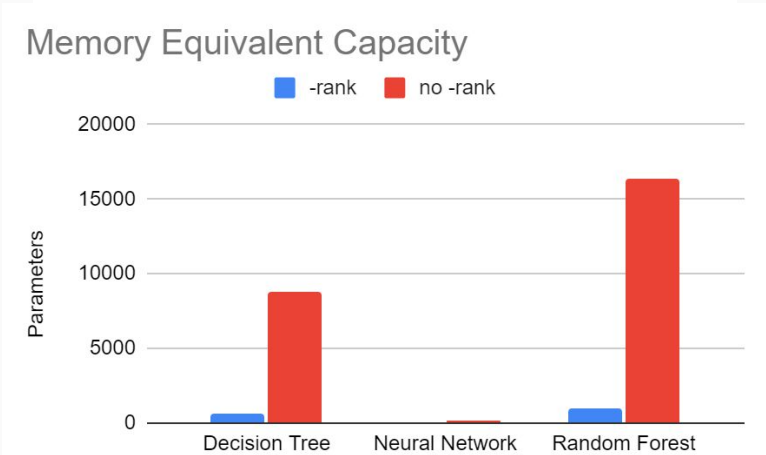
House Price Classification

- Goal: predict the sale price of a house
- Example Attributes: # of bedrooms, square feet
- Number of instances: 21613
- Number of attributes: 7
- Number of classes: 2
- Class balance: 58.11% 41.89%

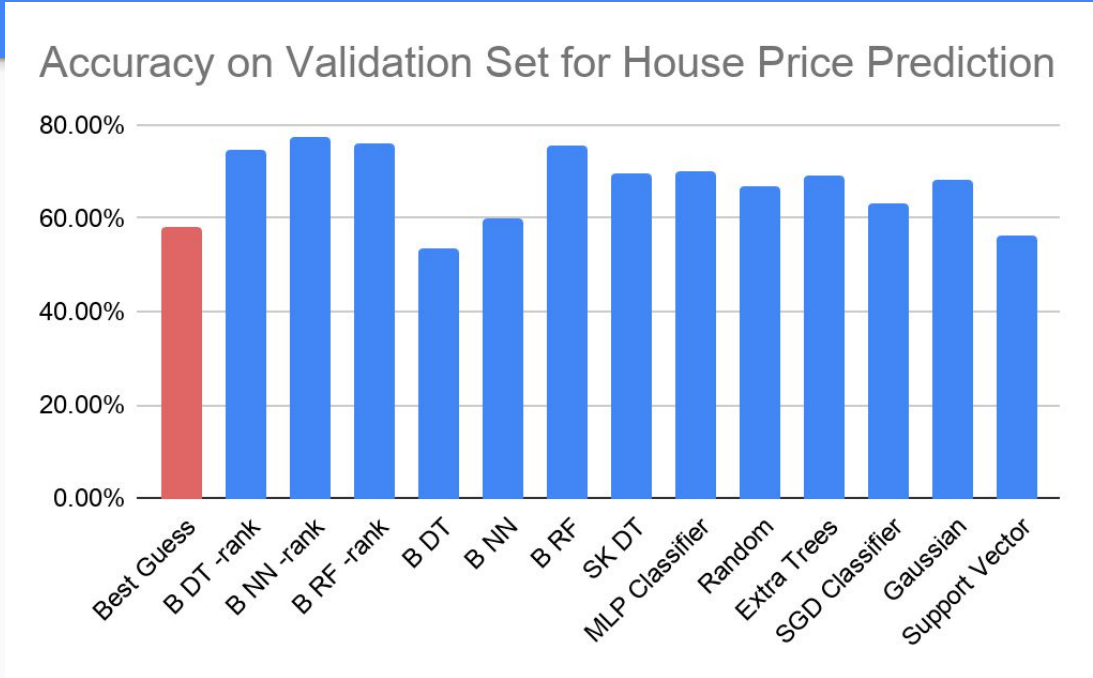
House Price Classification Measurements

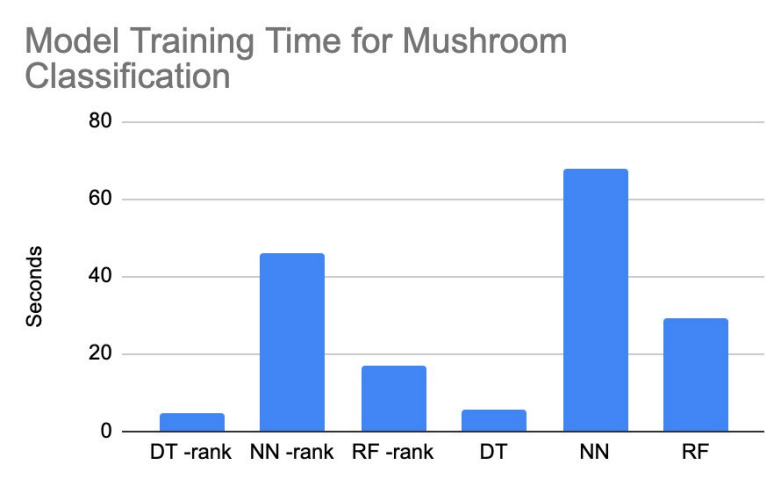






House Price Classification Model Results





House Price Classification Summary

Best Model:

- Brainome NN -rank

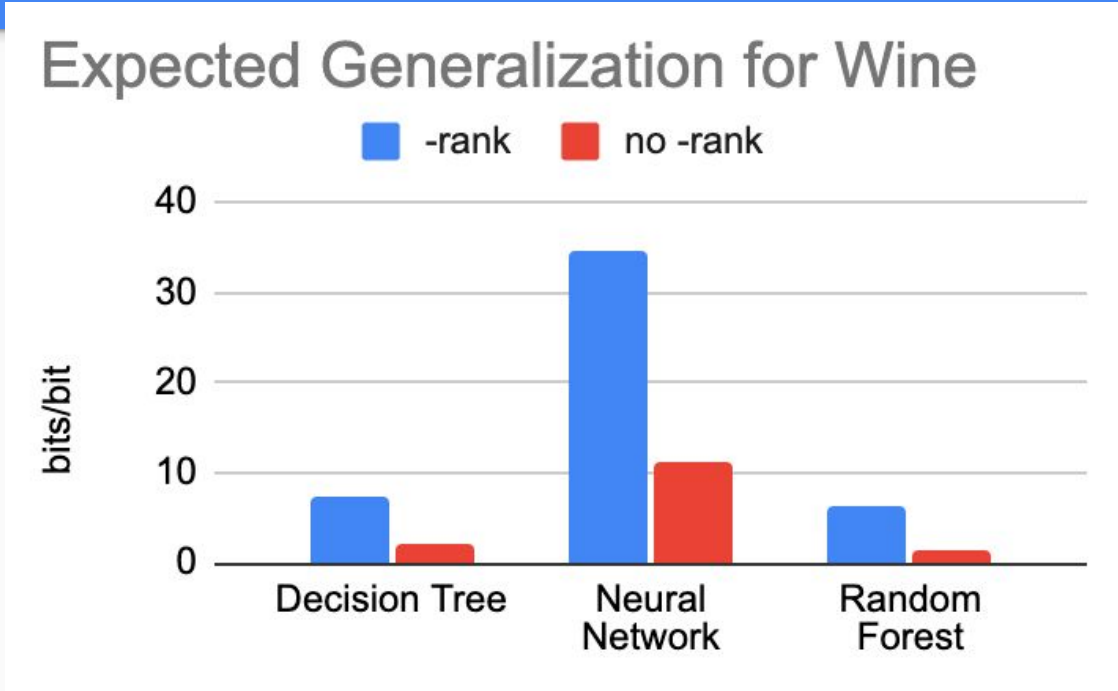
Useful measurement:

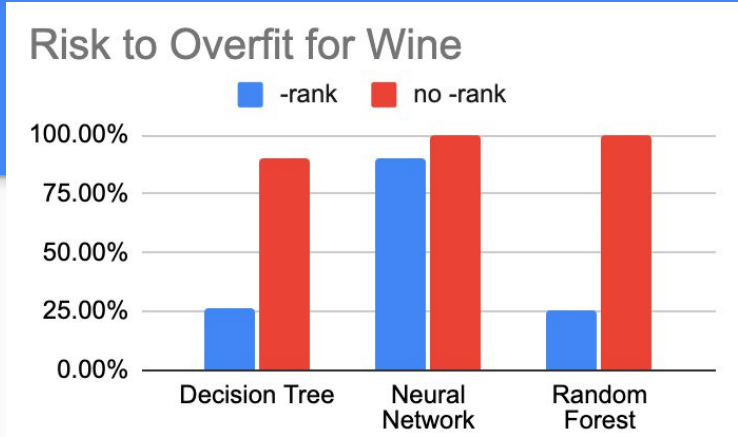
- Highest Generalization

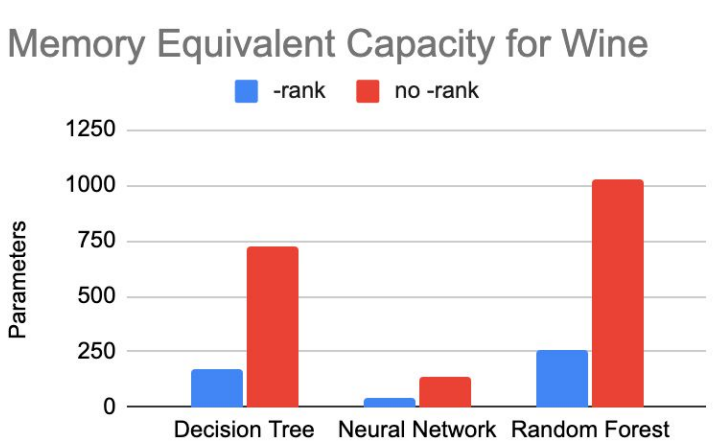
Wine Classification

- Goal: Predict the rating of a wine
- Example Attributes: alcohol, pH
- Number of instances: 1599
- Number of attributes: 12
- Number of classes: 2
- Class balance: 46.53% 53.47%

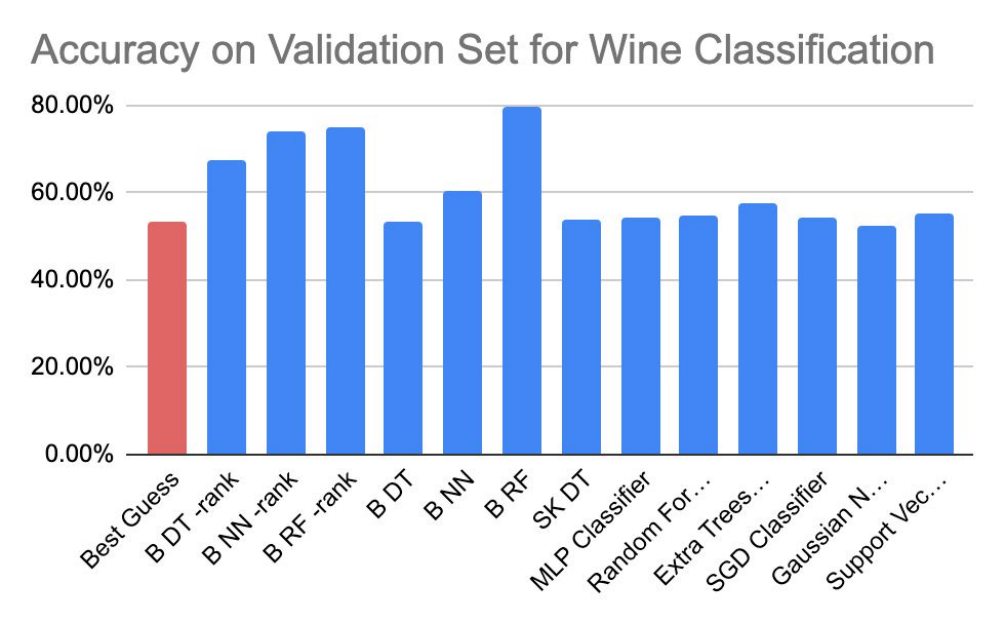
Wine Classification Measurements

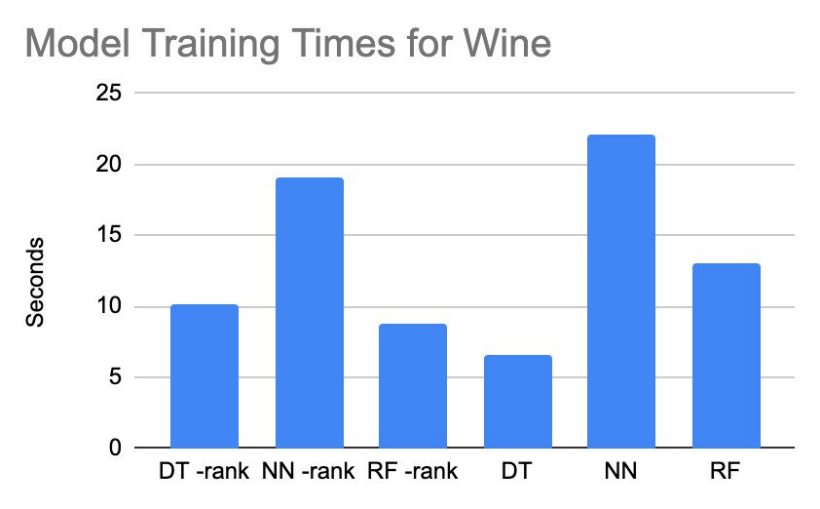






Wine Classification Model Results





Wine Classification Summary

Best Model:

- Brainome RF
- Brainome NN -rank second

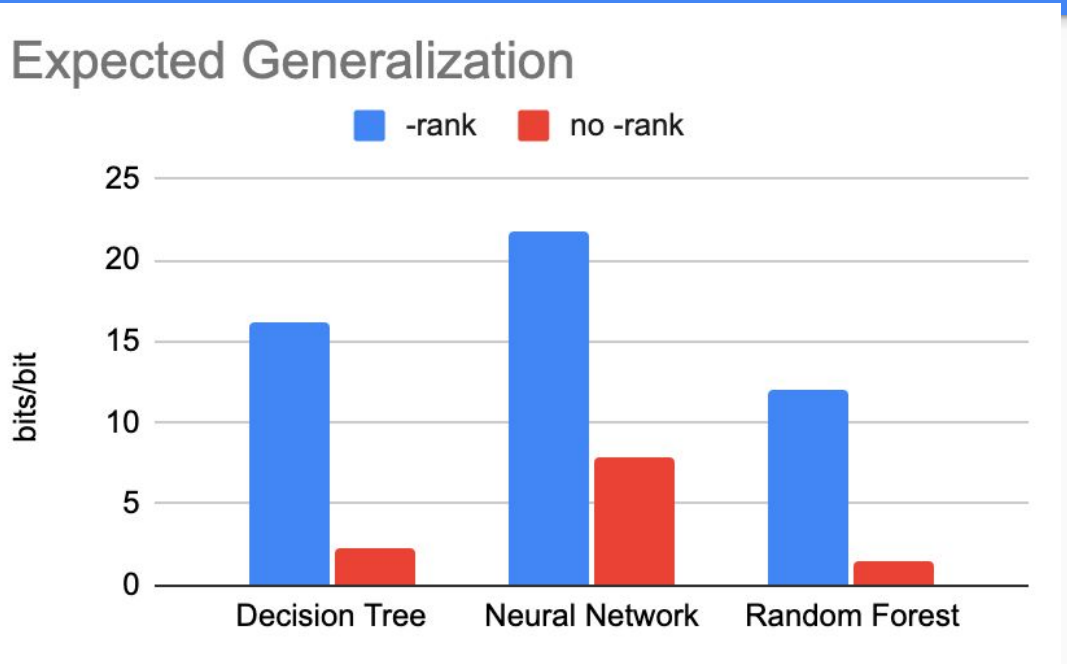
Useful measurement:

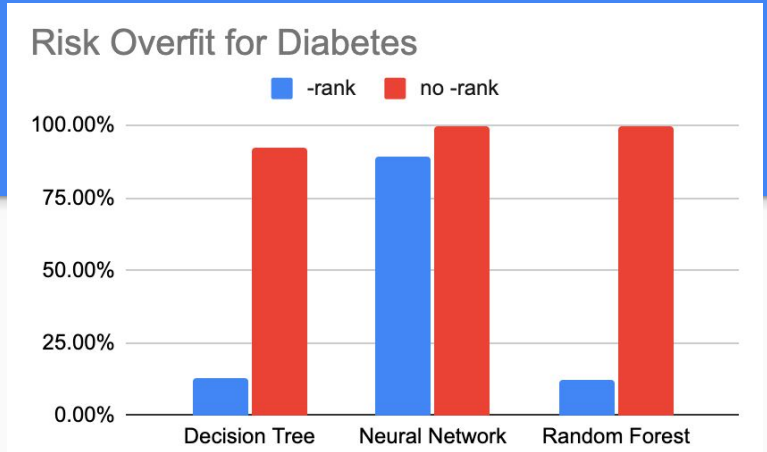
- Highest Generalization

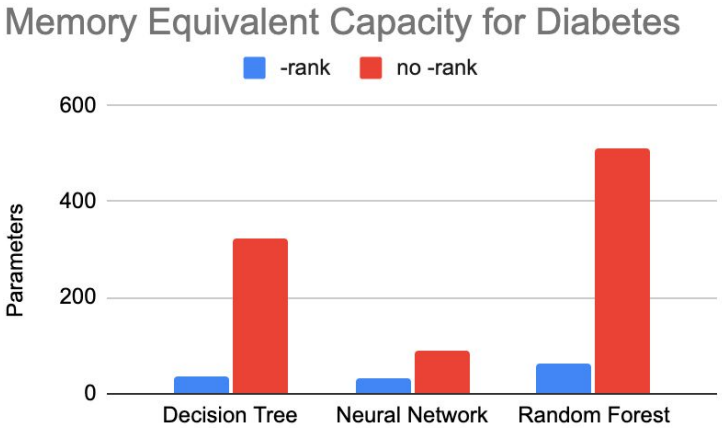
Diabetes Prediction

- Goal: predict if a person has diabetes
- Example attributes: glucose, BMI
- Number of instances: 768
- Number of attributes: 8
- Number of classes: 2
- Class balance: 65.1% 34.9%

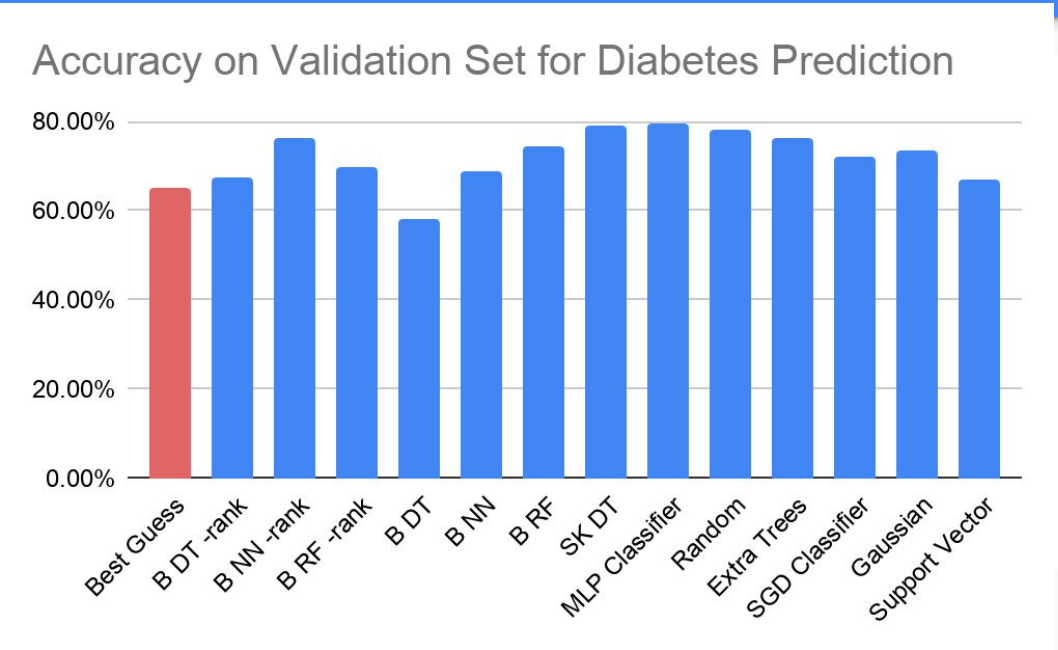
Diabetes Prediction Measurements

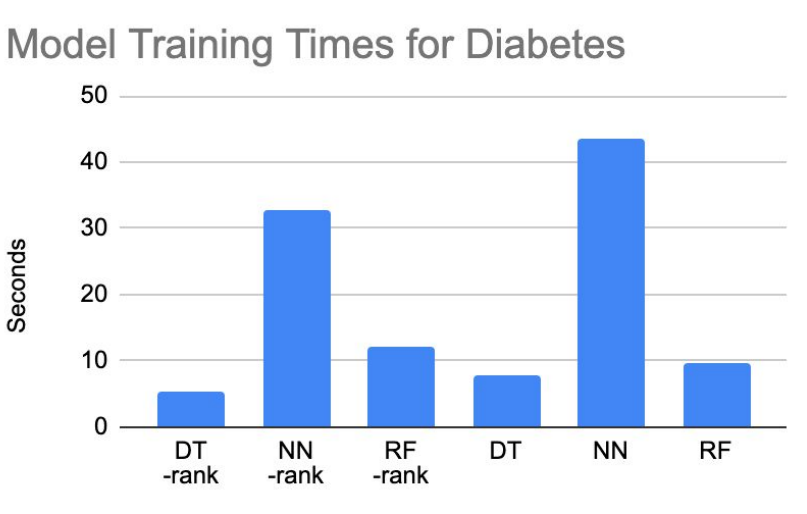






Diabetes Prediction Model Results





*Scikit Learn accuracies are after using Brainome Measurements to maximally improve models

Diabetes Prediction Summary

Best Models:

- Scikit - Learn MLP Classifier (NN)
- Brainome NN -rank

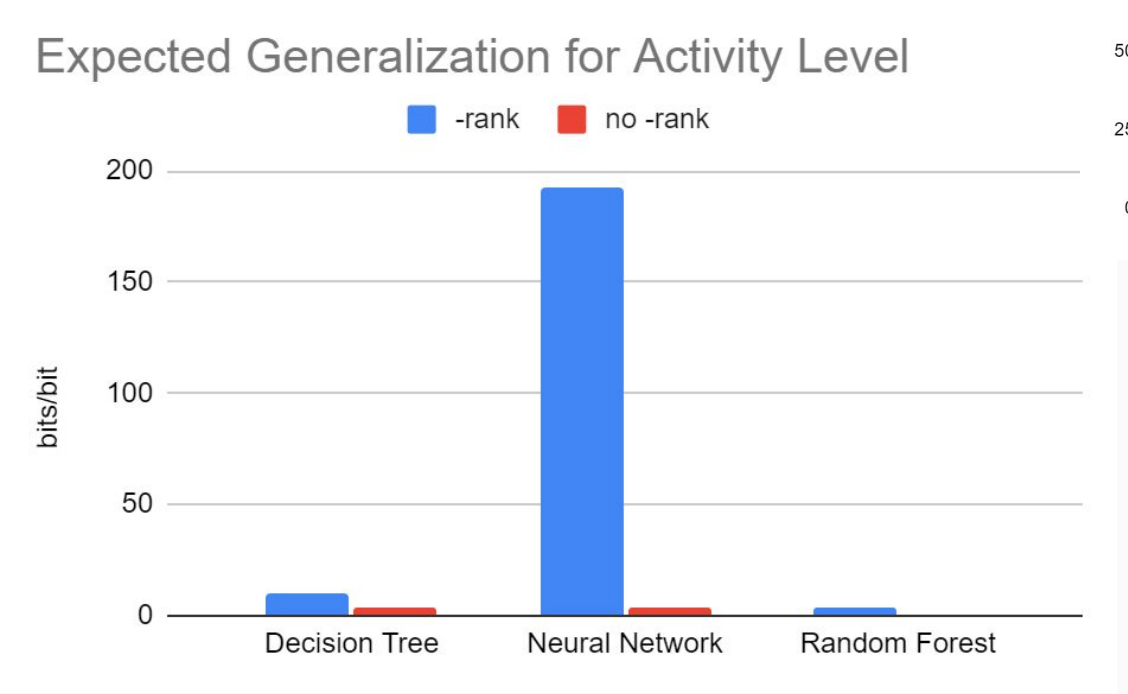
Useful measurement:

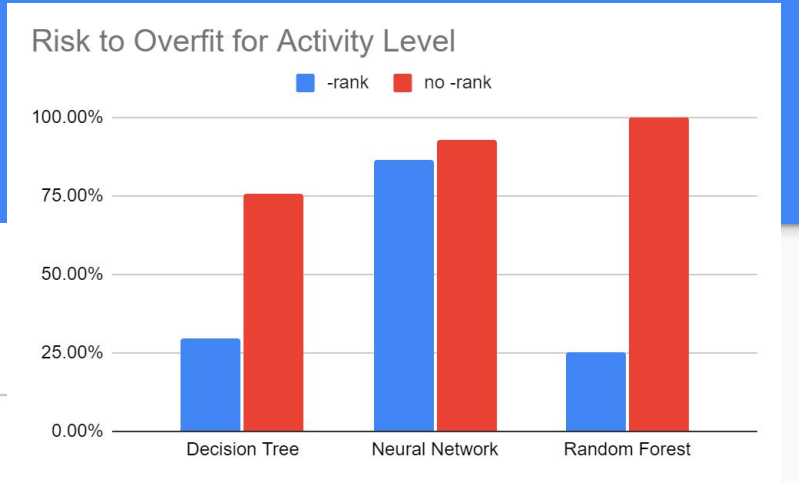
- Highest Generalization

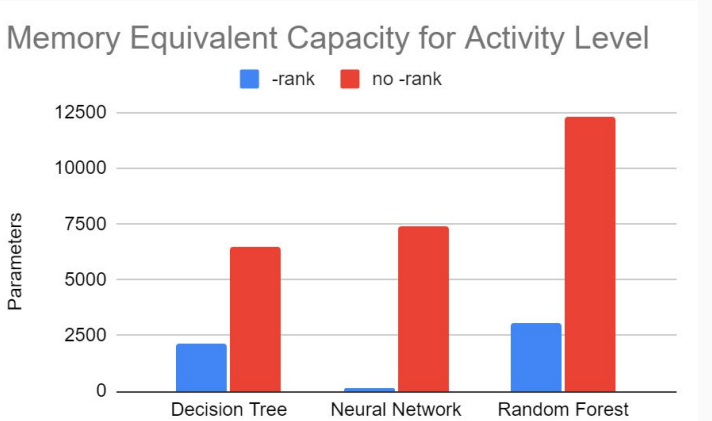
Activity Level Classification

- Goal: classify a person's activity level
- All attributes were preprocessed
- Number of instances: 10299
- Number of attributes: 561
- Number of classes: 6
- Class balance: 18.51% 17.25% 18.88% 16.72% 13.65% 14.99%

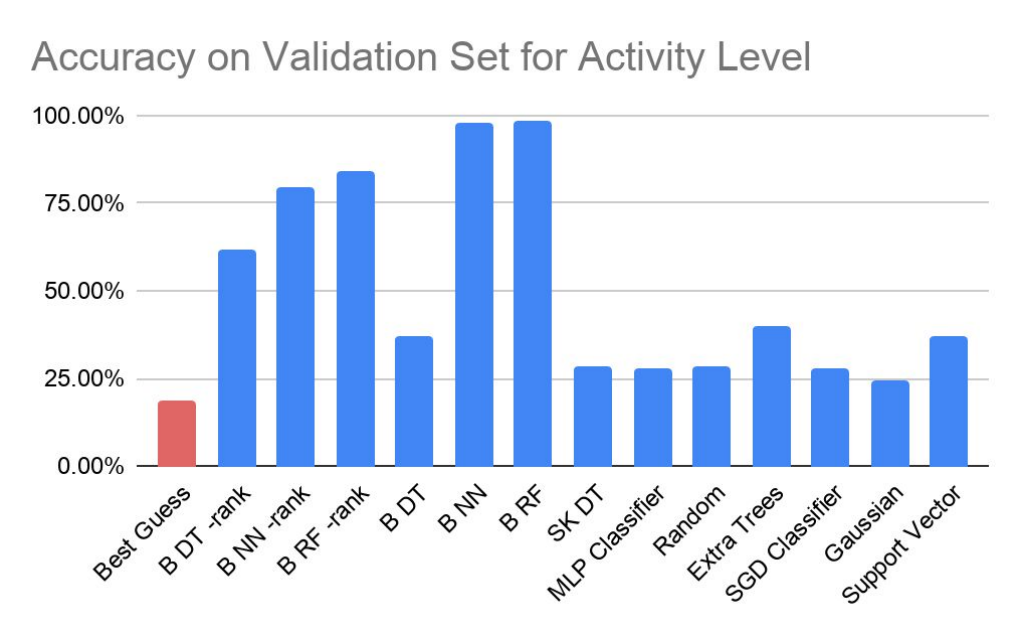
Activity Level Classification Measurements

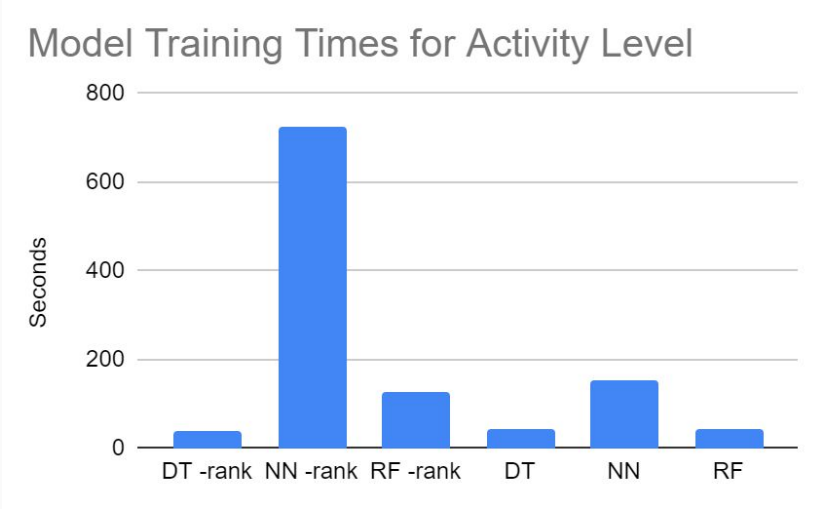






Activity Level Classification Model Results





Activity Level Classification Summary

Best Model:

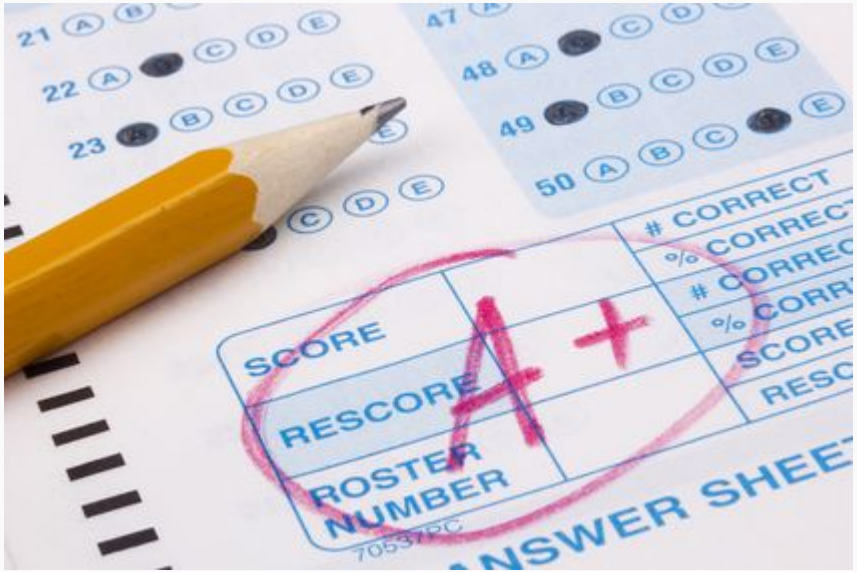
- Brainome RF
- Brainome NN

Useful measurement:

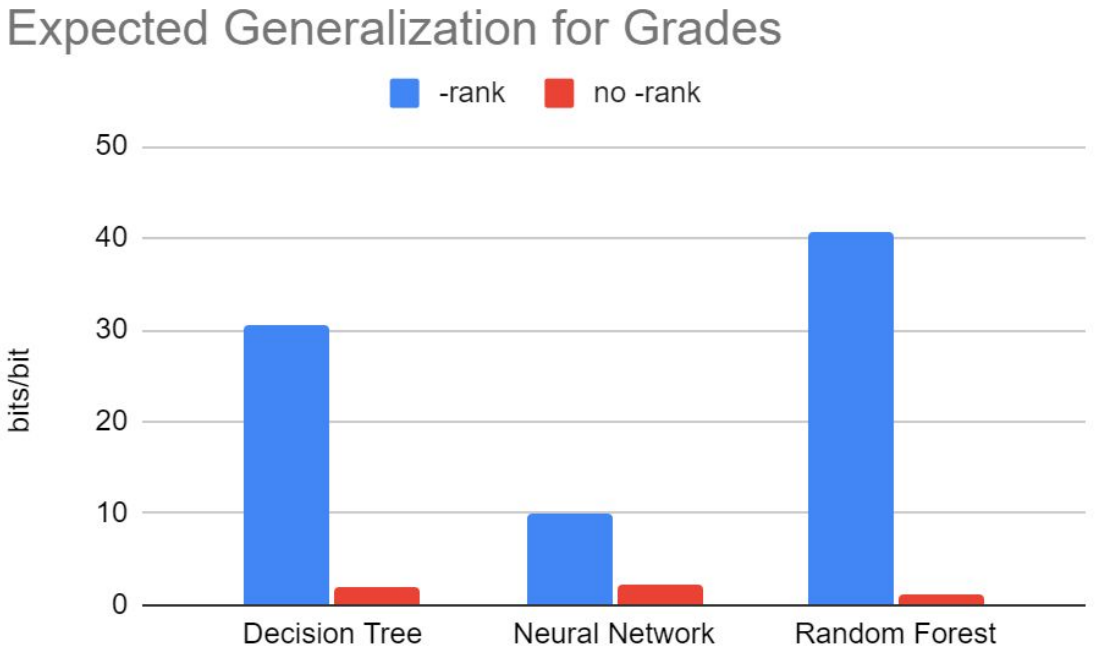
- Using alpha version of BTC, so measurement for RF are inaccurate
- NN does follow trend for measurements

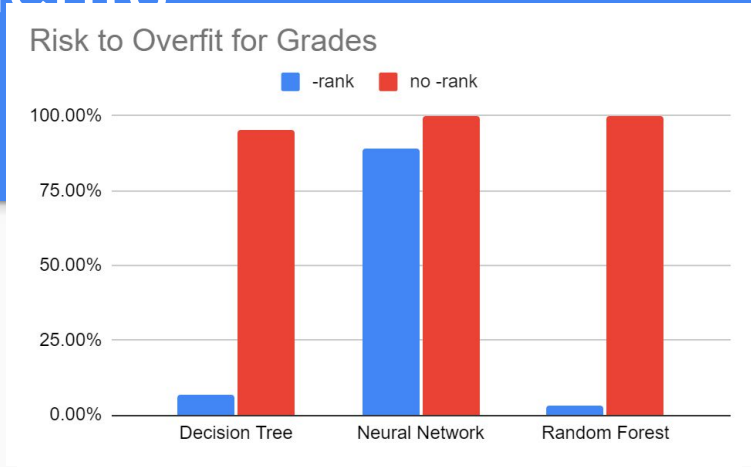
Grade Prediction

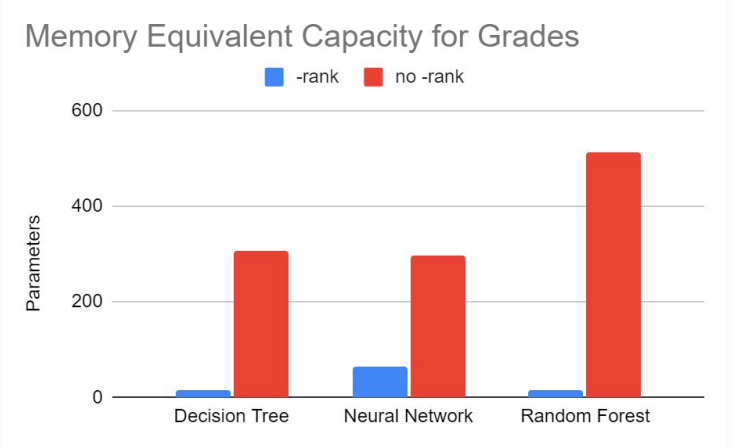
- Goal: Predict a students performance
- Example Attributes: study time, guardian's job
- Number of instances: 649
- Number of attributes: 31
- Number of classes: 2
- Class balance: 46.38% 53.62%



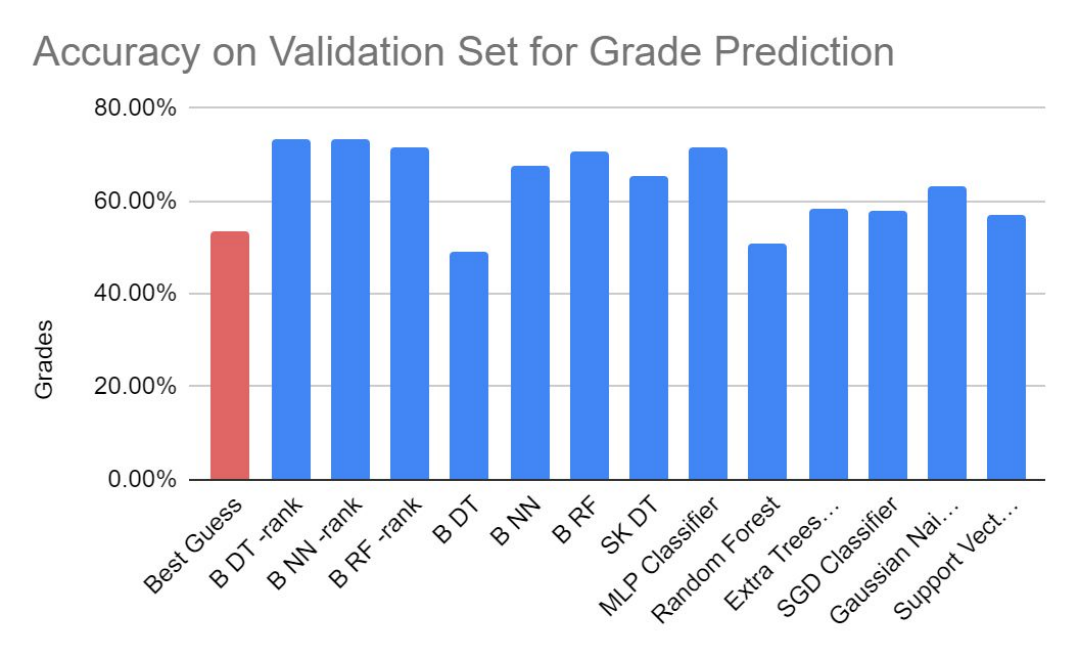
Grade Prediction Measurements

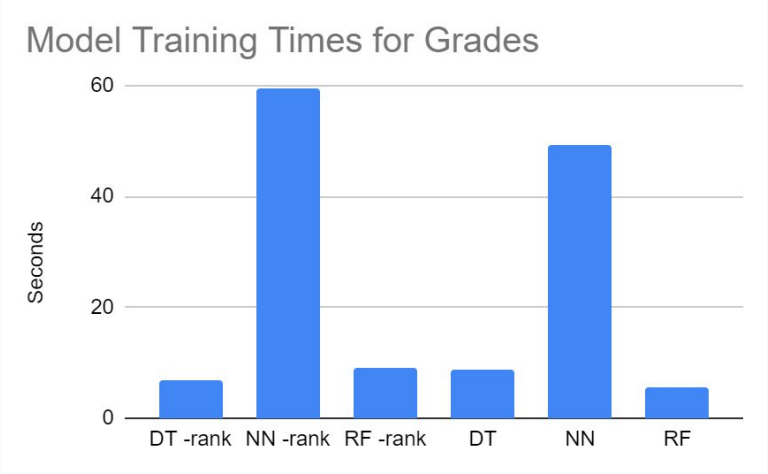






Grade Prediction Model Results





Grade Prediction Summary

Best Models:

- Brainome DT and RF -rank
- Brainome NN -rank

Useful measurement:

- DT and RF: Generalization
- NN: measurements are approximated

Objective Summary

Average Test Accuracy Improvement Over Best Guess

	Scikit Learn	BTC
Across All Models	6.49%	16.25%
Across Optimal Model for Each Dataset	12.80%	25.69%

*Scikit Learn accuracies are after using Brainome Measurements to maximally improve models

Any Questions?