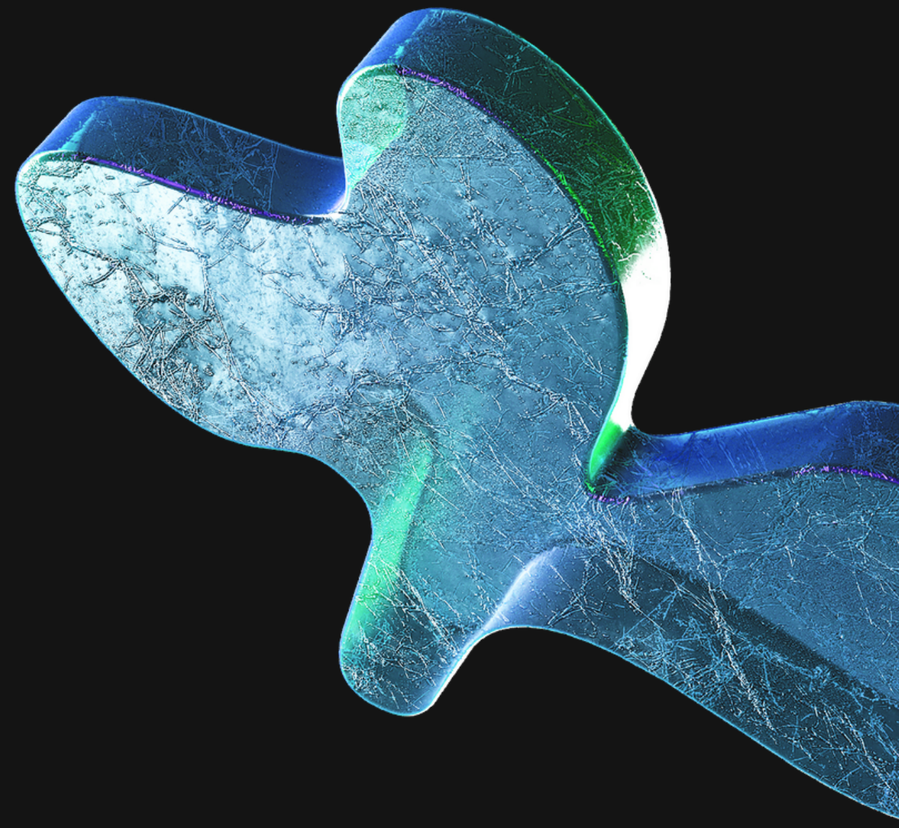
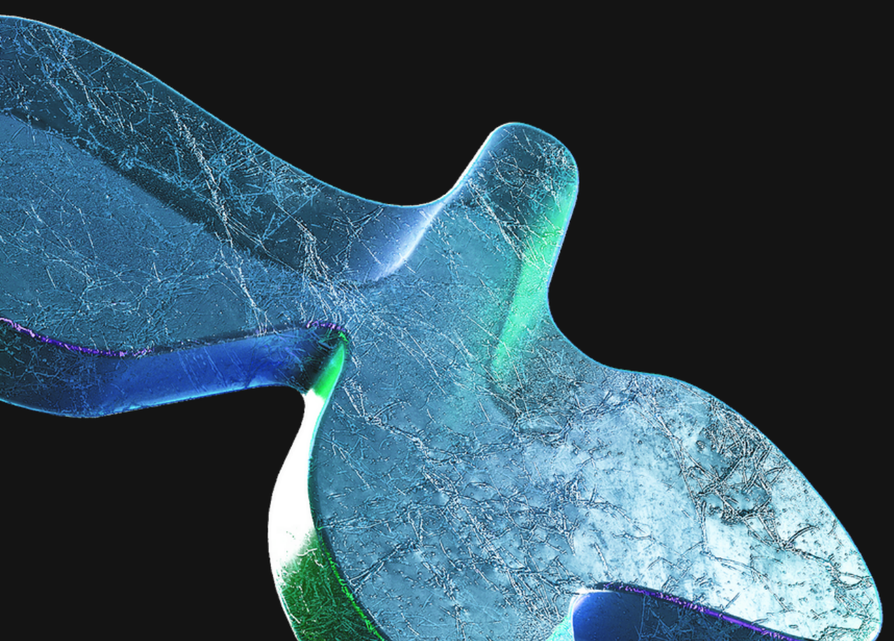
Options Pricing



Diana Wu Chen, Manhui Zhu, Hsiao-Yun (Angel) Tsui

DSO 530: Group Project
4/19/2024

Agenda



Introduction: Options

Value Prediction（Regression Models）

BS prediction（Classification Models）

Most Accurate Approach

Key Takeaways

Q&A

What is an Option?

Options are financial derivatives that give buyers the right, but not the obligation, to buy or sell an underlying asset at an agreed-upon price and date.

Call Option (Buy)

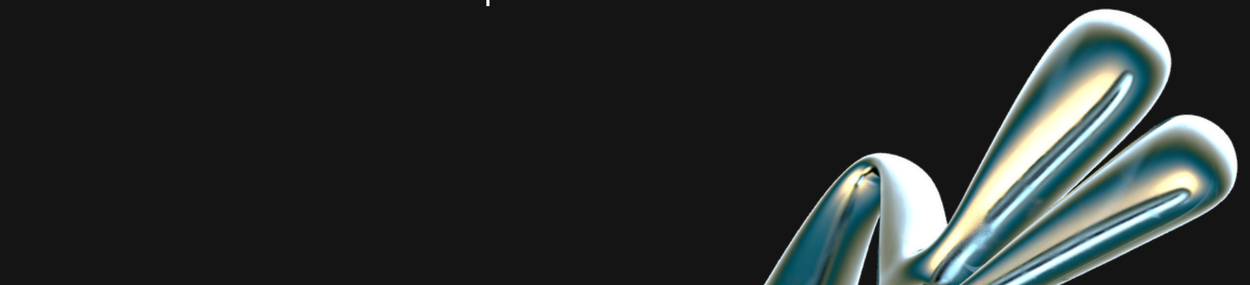
The right to buy a security at the specified price up to the date of expiration

- Not obligated to execute the option
- Gains are unlimited
- Investors look for a price rise

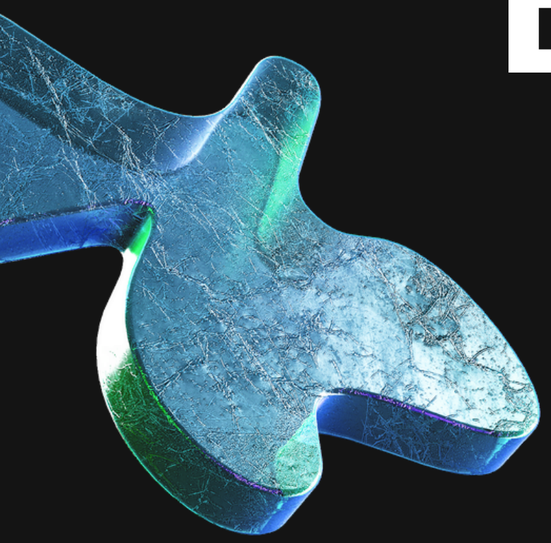
Put Option (Sell)

The right to sell a security at the specified price up to the date of expiration

- Not obligated to execute the contract
- Gains are limited
- Investors look for a fall in price



The Black-Scholes formula



The Black-Scholes formula:

$$C\_pred = S\Phi(d_1) - Ke^{-r\,\tau}\Phi(d_2)$$

$$d_1 = \frac{\log(S/K) + (r + \sigma^2)\tau}{\sigma\sqrt{\tau}}$$

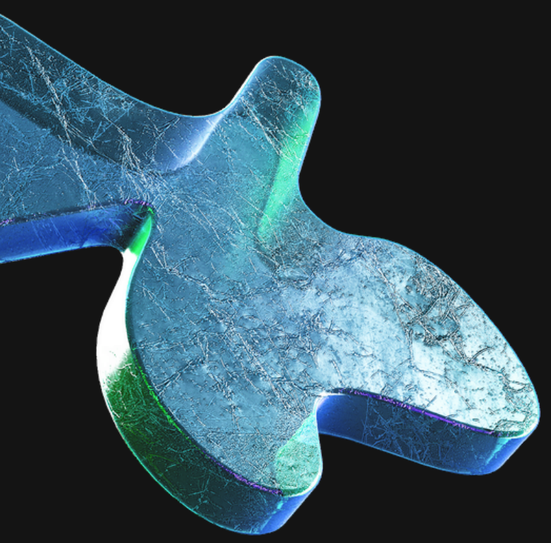
$$d_2 = d_1 - \sigma\sqrt{\tau}$$

The Black-Scholes equation requires five variables:

- Value (C): Current option value
- S: Current asset value
- K: Strike price of option
- r: Annual interest rate
- tau: Time to maturity (in years)

In our dataset: one more variables

- BS: Over (1), Under (0)

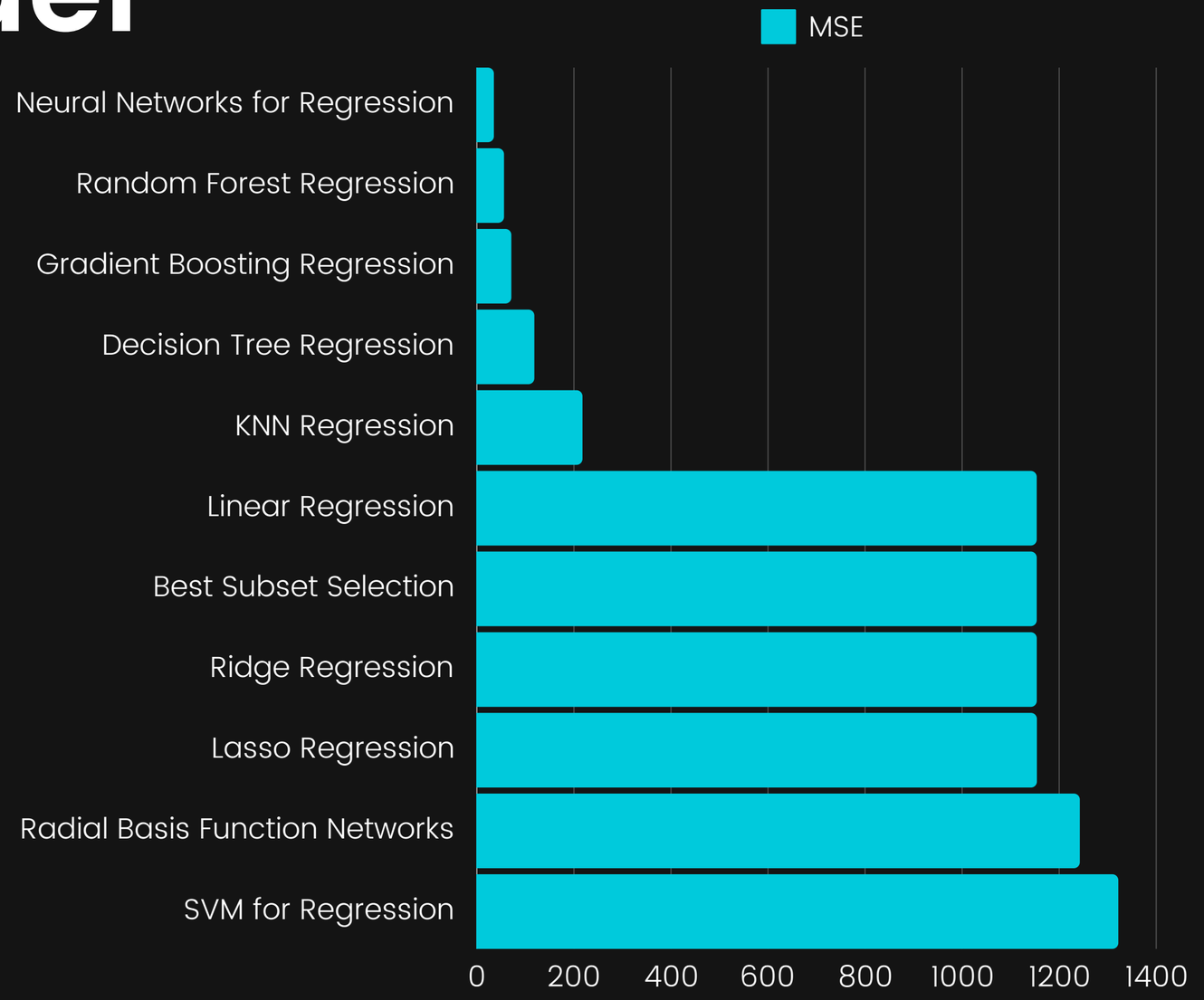


Value Prediction

(Regression Models)

Regression Model Performance

- Evaluation Matrix: Mean Squared Error (MSE)
- Scaling data first before fitting linear regression, KNN regression, Lasso & Ridge regression, and Neural Networks models, SVM, Gradient Boosting regression
- Top 3 models that give the smallest MSE:
  - 2-hidden layers Neural Networks
  - Gradient Boosting Regression
  - Random Forest Regression
- Neural Networks has the highest R^2: 0.9978 and lowest MSE: 35



Best Regression Model: Neural Networks

Input

Size = 4
4 Predictors

First Fully Connected Layer

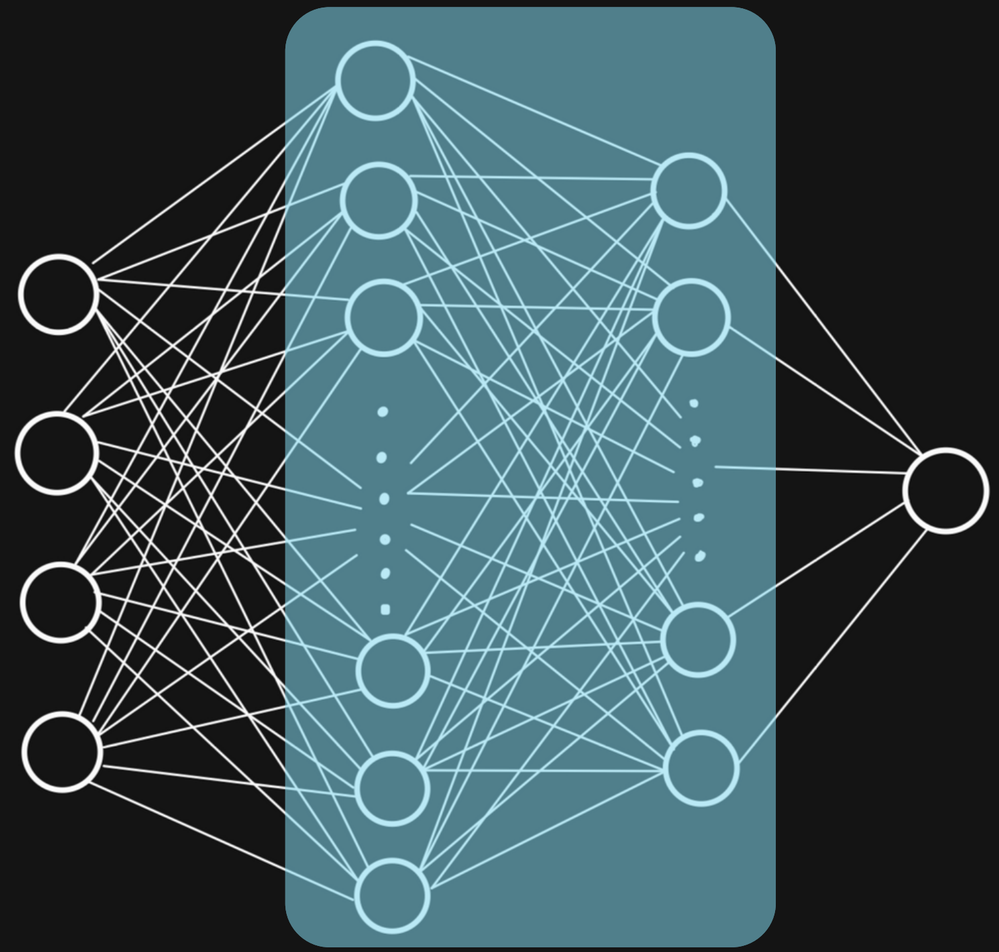
64 Hidden Neurons

Second Fully Connected Layer

32 Hidden Neurons

Output

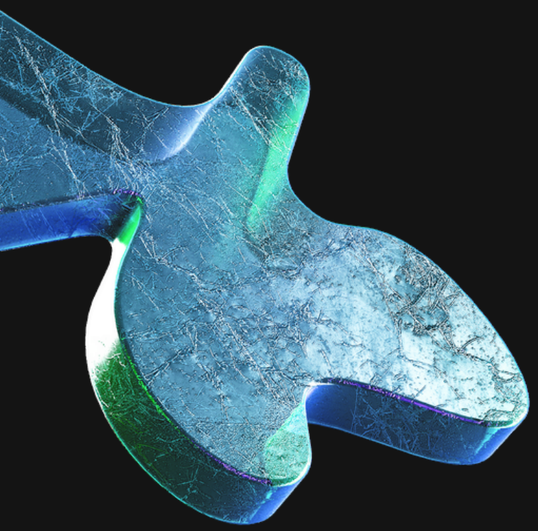
One output: ‘Value’
Activation: 'Linear'
Loss: 'mse'

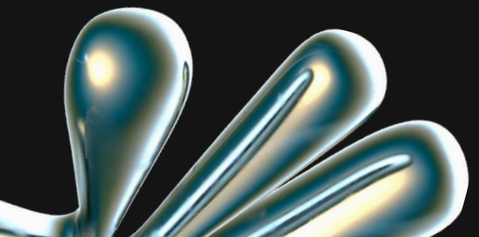


Input

2 Hidden Layers

Output



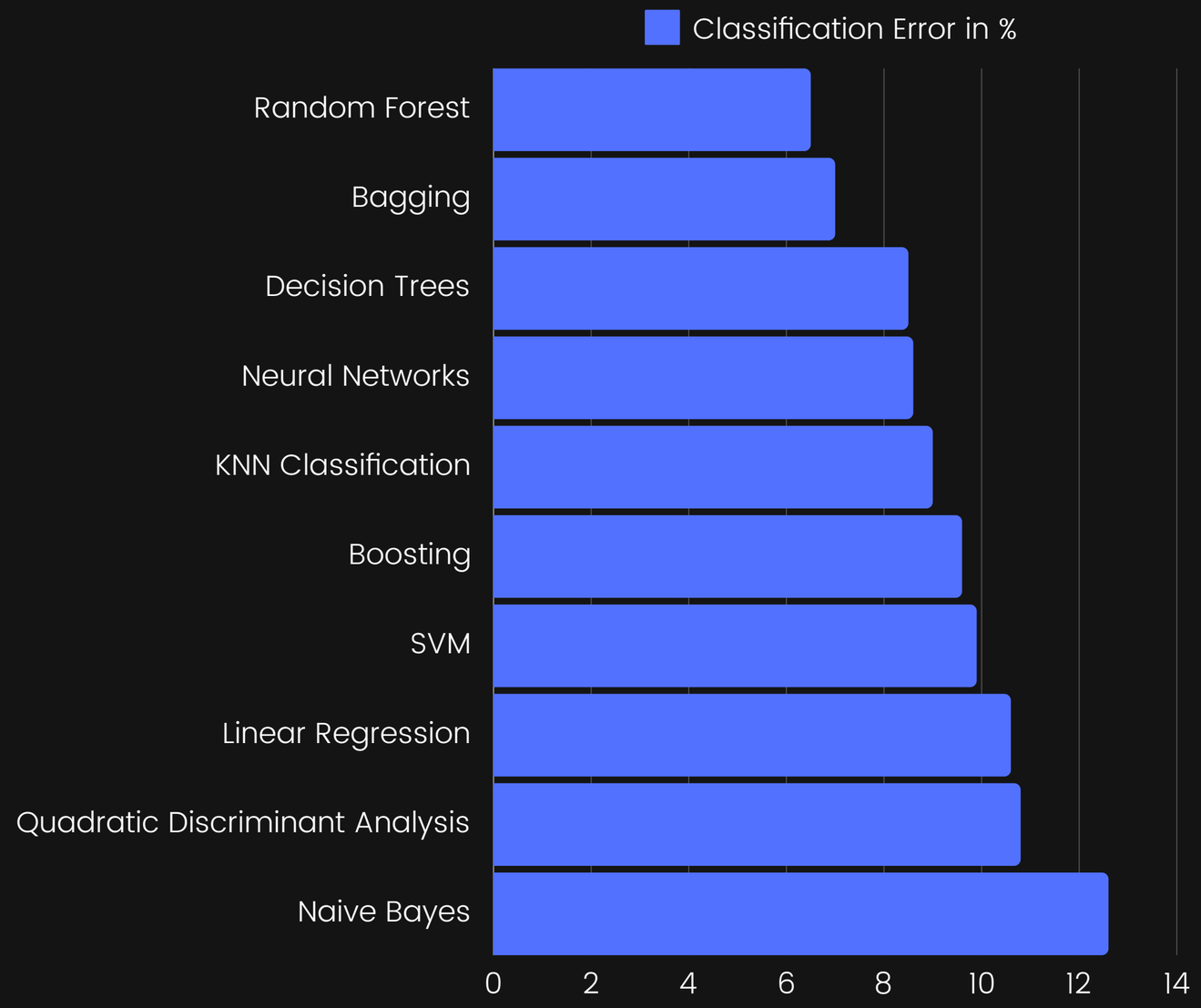


BS Prediction

(Classification Models)

Classification Model Performance

- Evaluation Matrix: Classification error
- Top 3 models that give the smallest classification error:
  - Random Forest
  - Bagging
  - Decision Tree Regression
- Random Forest has the lowest Classification Error of 6.9%

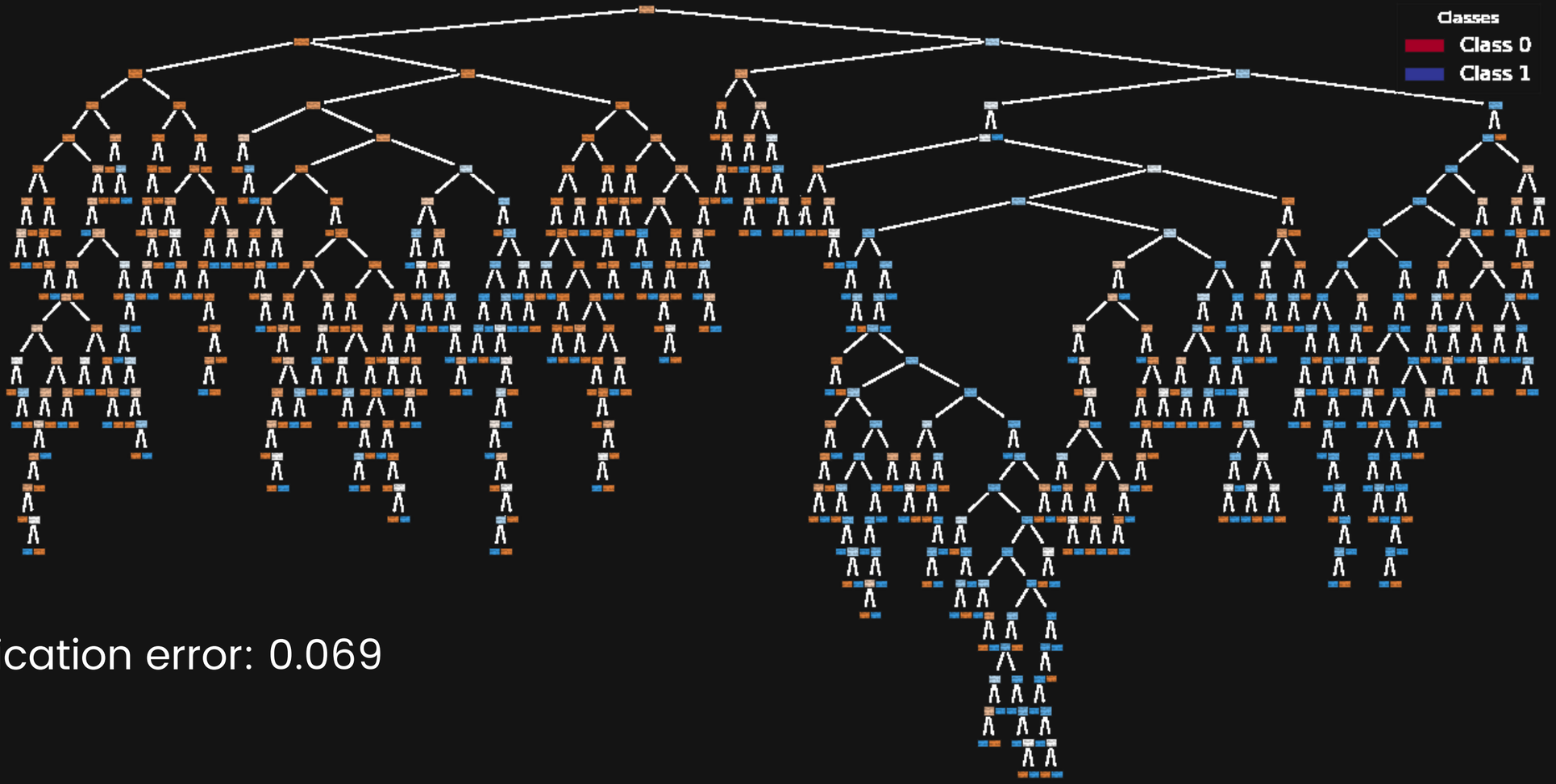


Best Classification Model: Random Forest

- **Feature Randomness:**
  - Helps in decorrelating the trees
- **Stronger Diversity:**
  - Reduces the likelihood of overfitting
- **Better Performance:**
  - Captures complex patterns in the data
- **Reduced Variance:**
  - Leads to more stable and reliable predictions
- **Efficiency:**
  - RF builds shallow trees, reducing computational burden of training and prediction, it can build deeper trees

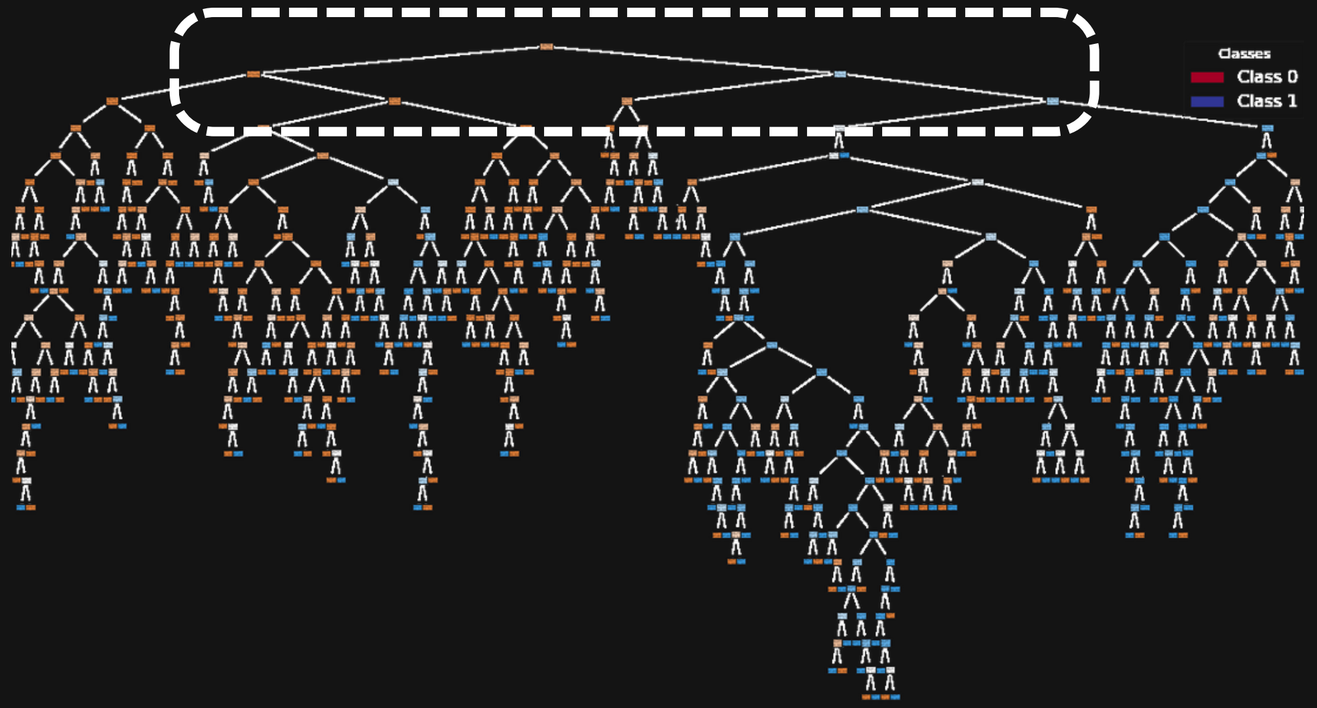
	Random Forest	Bagging
Classification Error	0.06999999995	0.06999999995

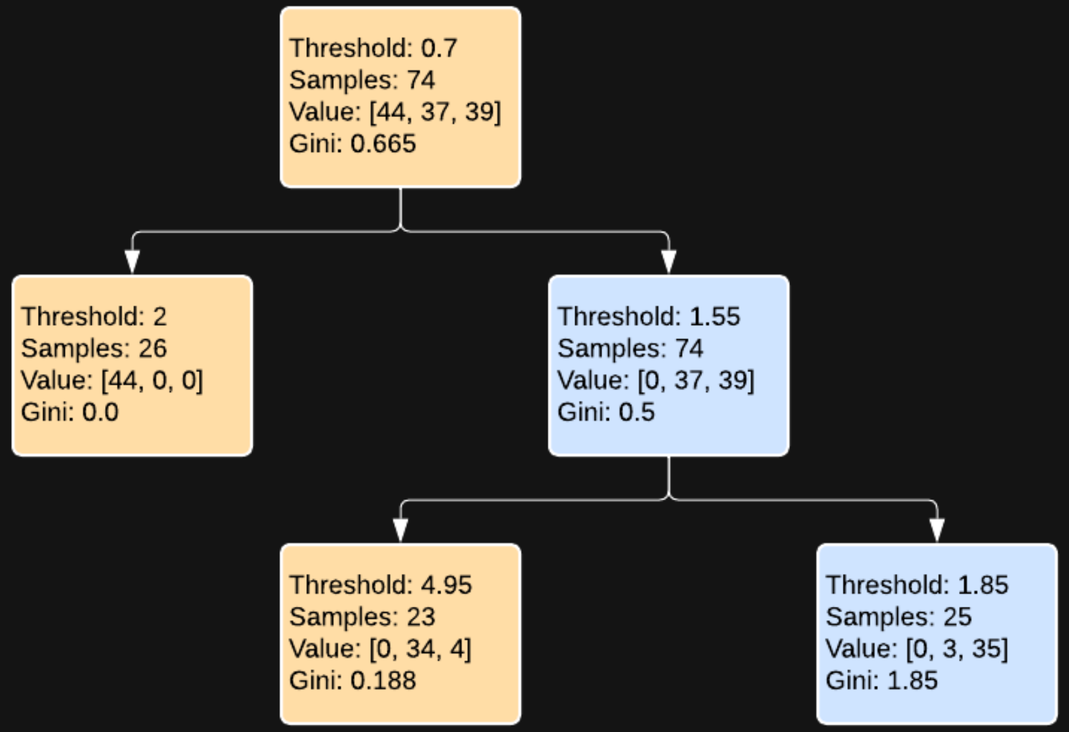
Best Classification Model: Random Forest



Classification error: 0.069

Best Classification Model: Random Forest

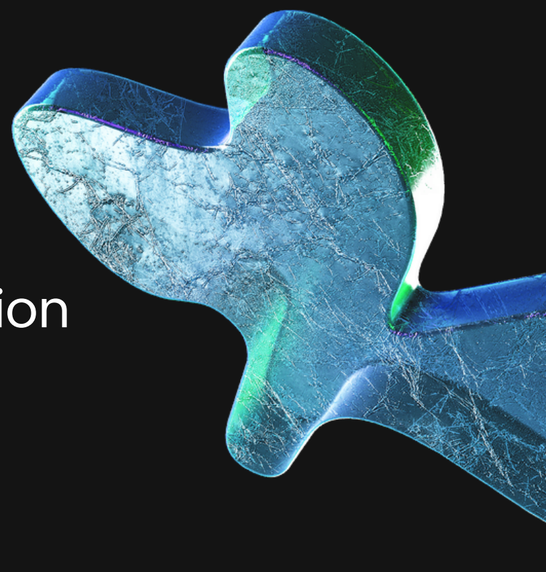




Classification error: 0.069

Key Takeaways

- Random Forest excels in both regression and classification models
  - Ensemble learning
  - Feature importance
  - Ability to handle missing data and outliers
  - Avoid overfitting
- Machine learning models outperform the Black-Scholes formula
  - Flexibility & Adaptability
    - Capture complex, non-linear relationships of data
    - Incorporate more features beyond the formula
  - Model Performance Improvement
    - Train and update the model with real-time data for prediction accuracy improvement



Thank You
(Q&A)