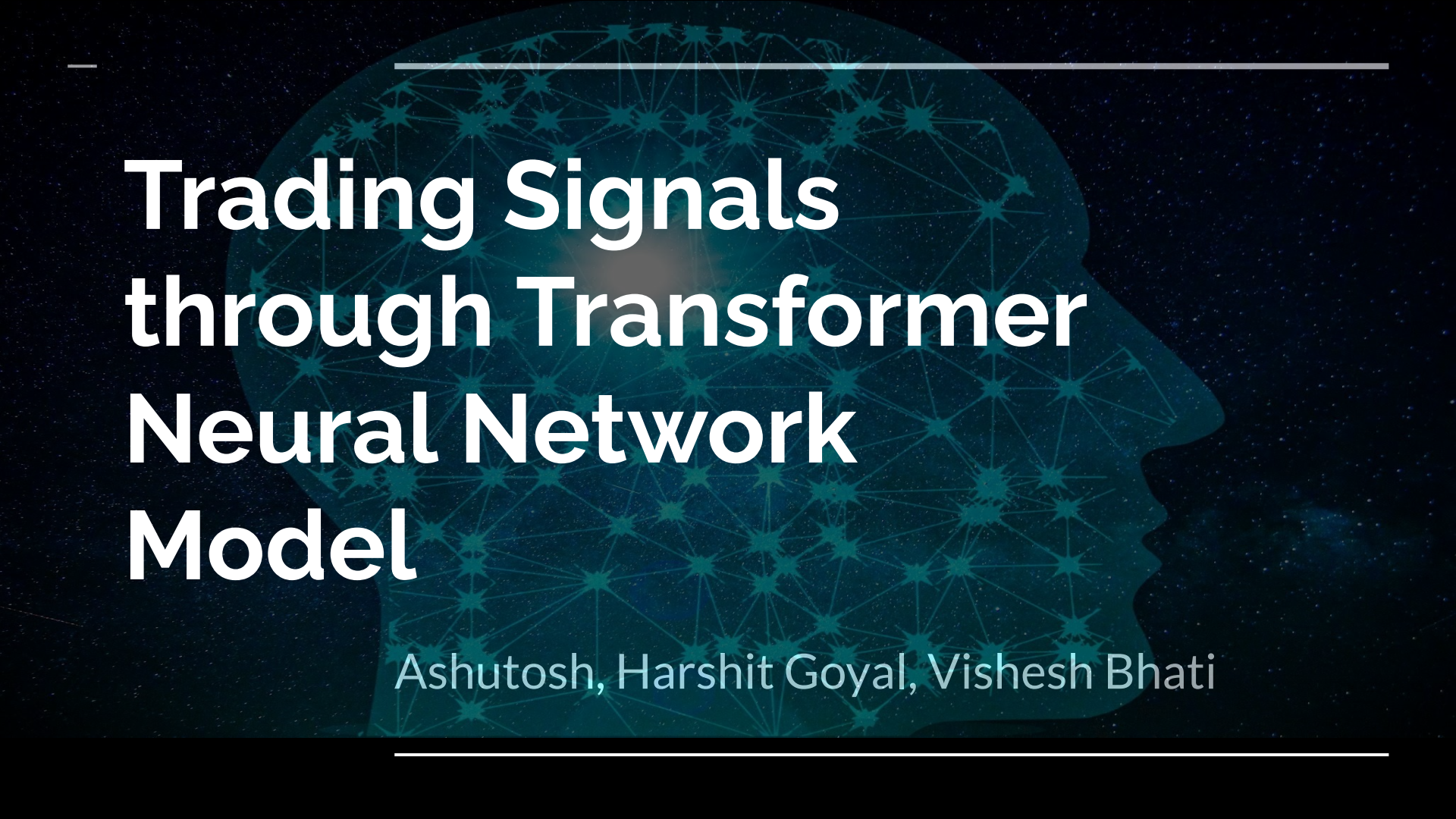

Trading Signals through Transformer Neural Network Model

Ashutosh, Harshit Goyal, Vishesh Bhati

Panel Data used for modelling:

Frequency : Daily

TimeFrame: 1992 to 2022

Data Size: 112 X 2.5 Mil. Records

Types of Parameters:

1. **Price Action: 6 parameters (Source: WRDS)**
2. **Financial Statement: 68 parameters from financial statement of each permno (CapitalIQ)**
3. **Technical Indicators: 36 parameters from 36 individual technical indicators. (Calculated)**

Feature Engineering

1. Breusch Pagan test- To find varying variance p-value - 0 (Reject the Null Hypothesis) H0: The residuals have constant variance
2. Huber White process- Data Transformation
3. Best fit Linear Regression model: Best Fit model for the data set was Elastic net - First attempt High R^2 ~ 1
4. Light GBM - To find significant features (Features reduced to 65 from 112)
5. R^2 still close to 1

Transformer Architecture

So, What's a Transformer Model?

- ➔ A transformer model is a neural network that learns context and thus meaning by tracking relationships in sequential data.
- ➔ The Transformer comprises an Encoder and a Decoder, each consisting of multiple layers. The Encoder, responsible for processing input data, is equipped with a stack of dense layers (Relu). On the other hand, the Decoder, responsible for generating trade signals, has linear activation function model.
- ➔ The rolling and recursive window strategy, coupled with the iterative training process, ensures that our model remains relevant and effective.
- ➔ We predict returns for each stock in a rolling window fashion and then calculate trading signals [-1, 1] based on previously set thresholds for each stock.

```
156/156 [==============================] - 1s 4ms/step
64/64 [==============================] - 1s 10ms/step - loss: 0.1751 - val_loss: 0.1329
156/156 [==============================] - 1s 4ms/step - loss: 0.0803
156/156 [==============================] - 1s 4ms/step
R2 from paper for PERMNO 10104.0: [[0.5699506]]
156/156 [==============================] - 1s 4ms/step
64/64 [==============================] - 1s 11ms/step - loss: 0.1202 - val_loss: 0.1013
156/156 [==============================] - 1s 4ms/step - loss: 0.0583
156/156 [==============================] - 1s 4ms/step
R2 from paper for PERMNO 10107.0: [[0.6765156]]
148/148 [==============================] - 1s 4ms/step
63/63 [==============================] - 1s 10ms/step - loss: 0.2062 - val_loss: 0.1132
148/148 [==============================] - 1s 4ms/step - loss: 0.0728
148/148 [==============================] - 1s 4ms/step
R2 from paper for PERMNO 10138.0: [[0.7211982]]
148/148 [==============================] - 1s 4ms/step
64/64 [==============================] - 1s 10ms/step - loss: 0.1525 - val_loss: 0.1022
148/148 [==============================] - 1s 4ms/step - loss: 0.0519
148/148 [==============================] - 1s 4ms/step
```

```
{10104.0: array([[[ 1]],

        [[ 1]],

        [[ 1]],

        ...,

        [[-1]],

        [[ 1]],

        [[ 1]]]),
 10107.0: array([[[ 1]],

        [[ 1]],

        [[ 1]],

        ...,

        [[-1]],

        [[ 1]],

        [[ 0]]]),
```