

Convolutional Neural Networks for Image Based Stock Buy/Sell Categorization



Ashay Deochand

Introduction

This project uses a 2-D Convolutional Neural Network based on image processing properties. In order to convert financial time series into 2-D images, 50 different technical indicators each with different parameter selections are utilized.

As a result, 50x50 sized 2-D images are constructed. Each image is then labeled as Buy, Sell or Hold depending on the hills and valleys of the original time series.

For analyzing and developing forecasting models from financial data, there are two main approaches: technical analysis and fundamental analysis.

Fundamental analysis can be implemented through examining company specific financial data such as balance sheet, cash flow, return on assets. Alternatively , technical analysis can be implemented through analyzing past financial time series data using mathematical and/or rule-based modeling

Project Objective

- The goal of the project: train a neural network to identify optimal points to buy or sell a stock using numerical financial indicators.
- CNN trained to recognize points of inflection where there is a high probability of price reversal.
- Three class labels: Buy / Sell / Hold
- 95 Largest companies listed on the Toronto Stock Exchange (TSX)
- 22 years of daily price data for time period from Jan-2000 till Dec-2022 (454 000 daily data points)

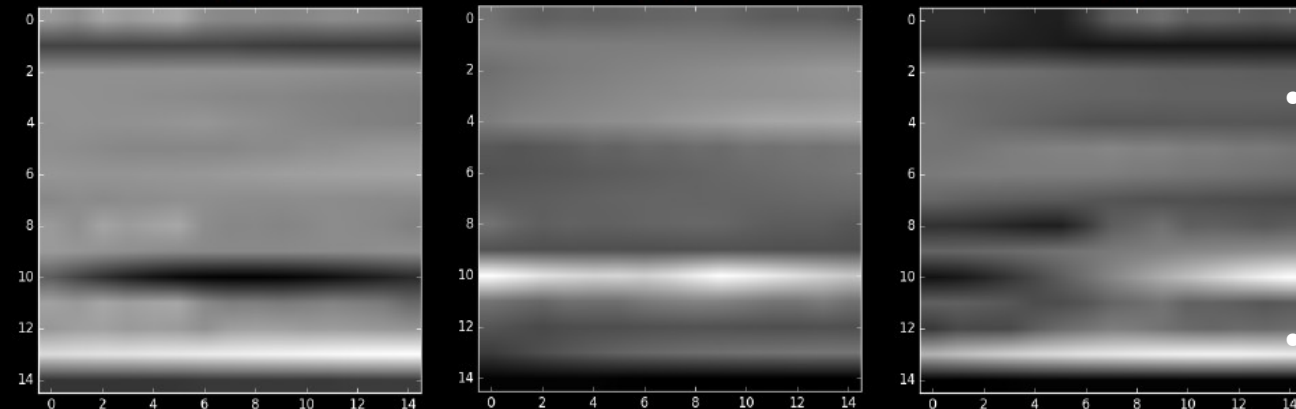


Target Class Labels

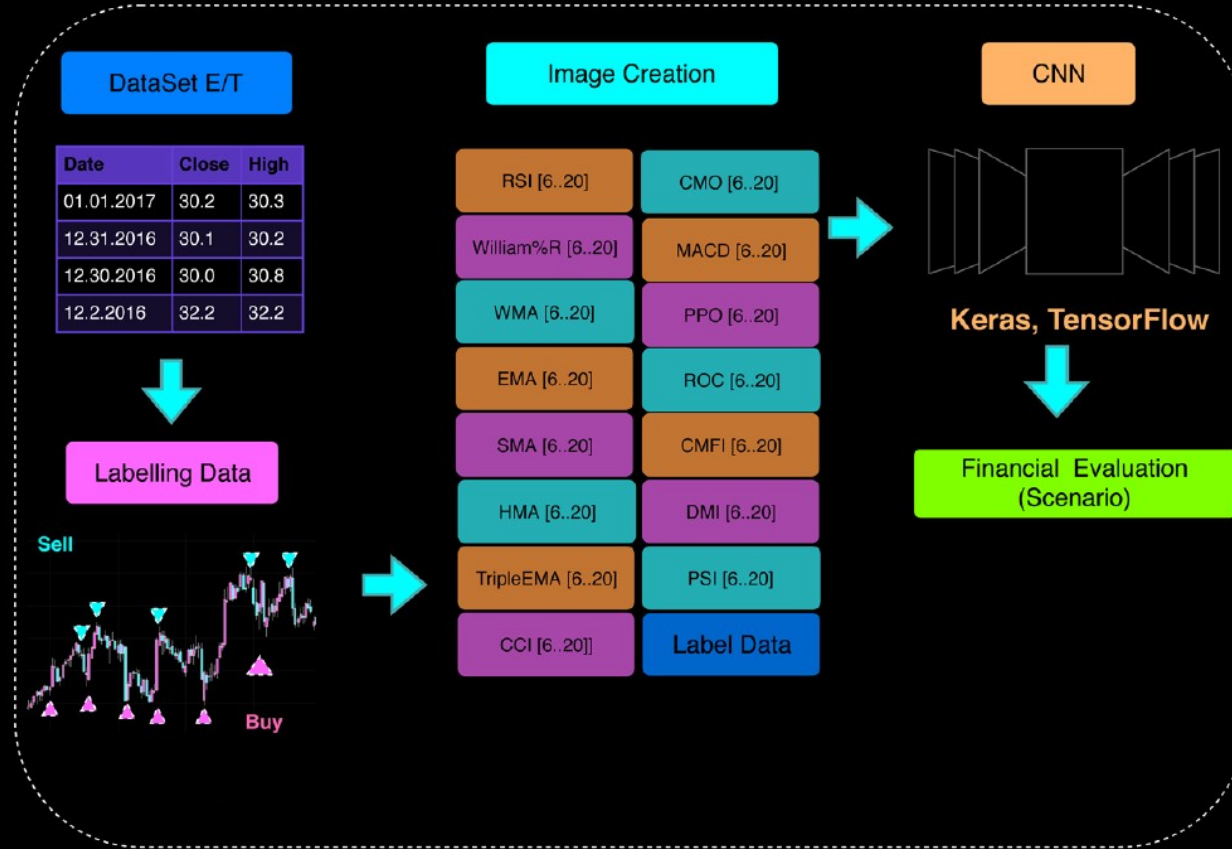
Royal Bank of Canada - Stock Price with Buy/Sell Labels



- Daily close prices are marked as “Sell”, “Buy”, or “Hold” by determining the top and bottom points in a sliding period window.
- One day before and after labeled together to increase Buy/Sell training size
- Circa. 50,000 total images.
18,000 Buy images; 18,000 Sell images and 19,000 Hold images.



Feature Engineering



Indicator Groups

Overlap Studies

Momentum Indicators

Volume Indicators

Volatility Indicators

Price Transform

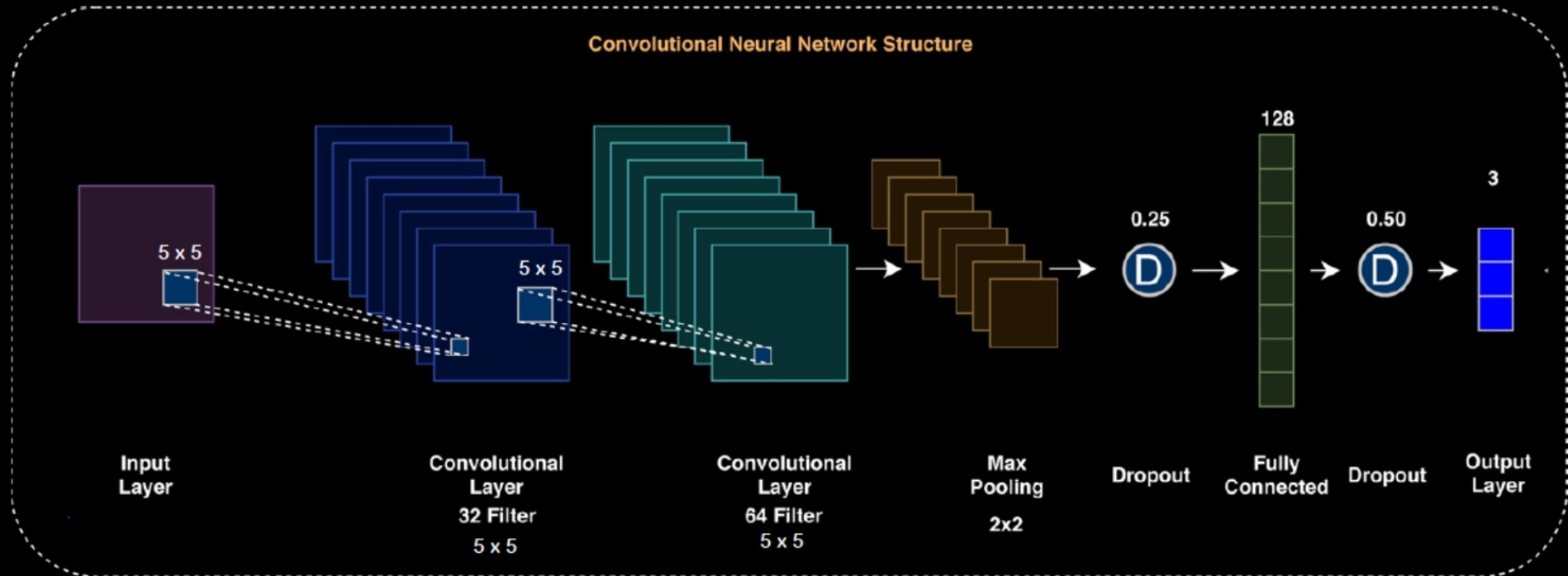
Cycle Indicators

Pattern Recognition

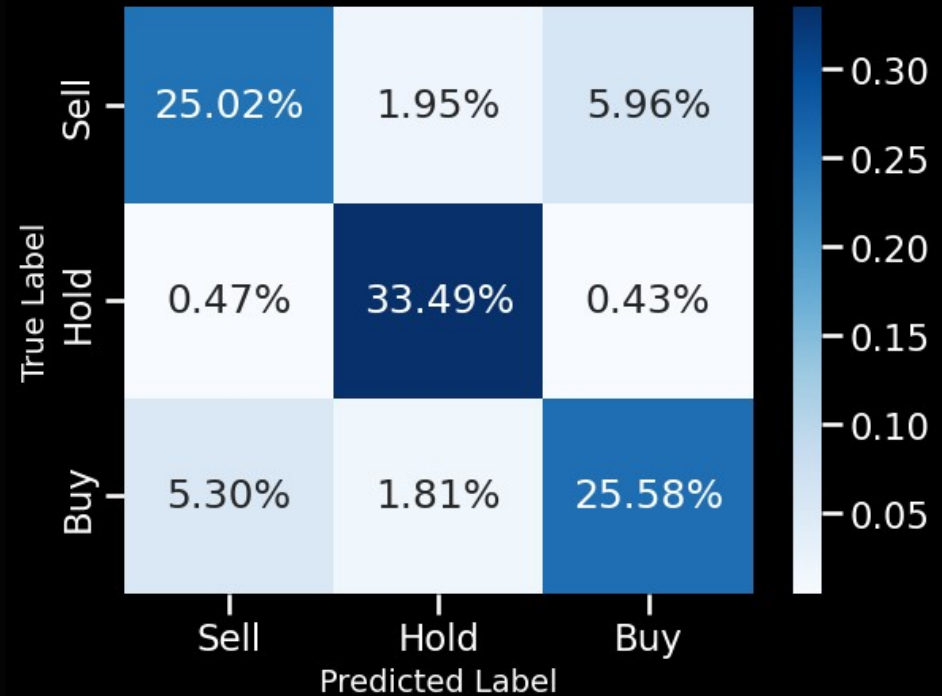
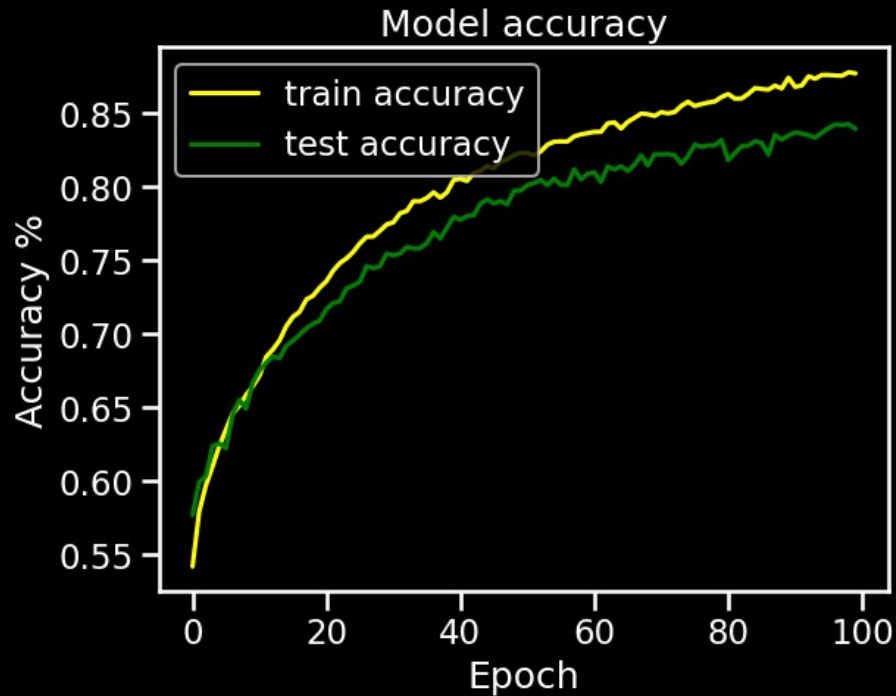
Historic Log Returns

Model Architecture

Input layer (50x50), two convolutional layers (50x50x32, 50x50x64), a max pooling (2x2x64), two dropout (0.25, 0.50), fully connected layers (128), and an output layer (3). Dropout layers are added to prevent overfitting.



Results: 80/20 Hold Out Validation

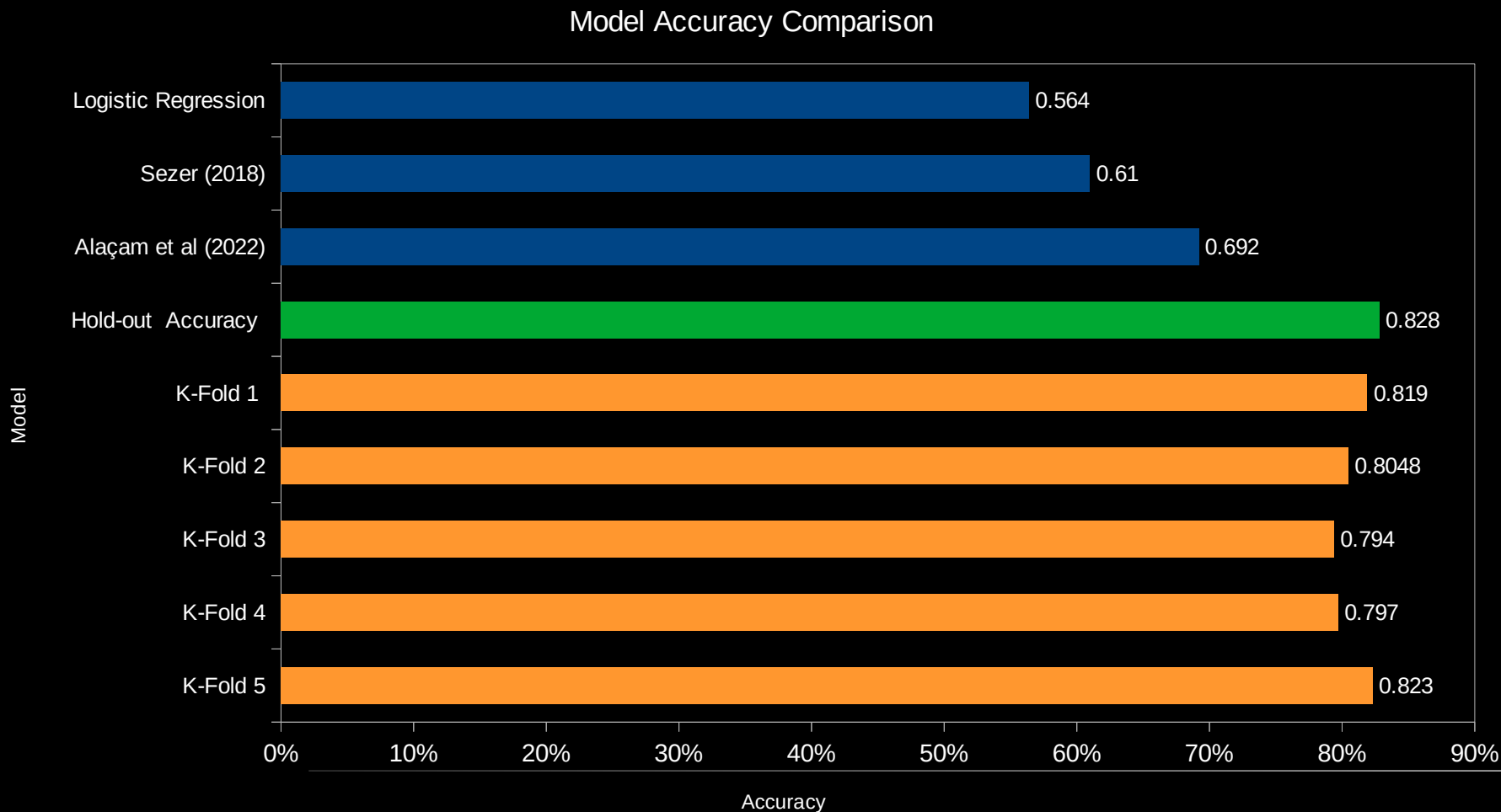


Test Accuracy: 82.87%

F1 score: 0.85

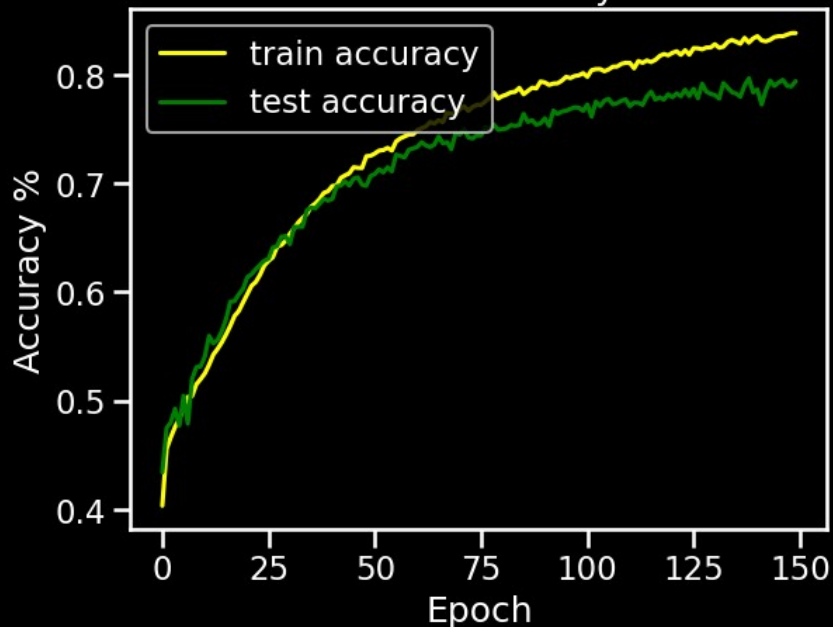


Results: K-Fold Validation

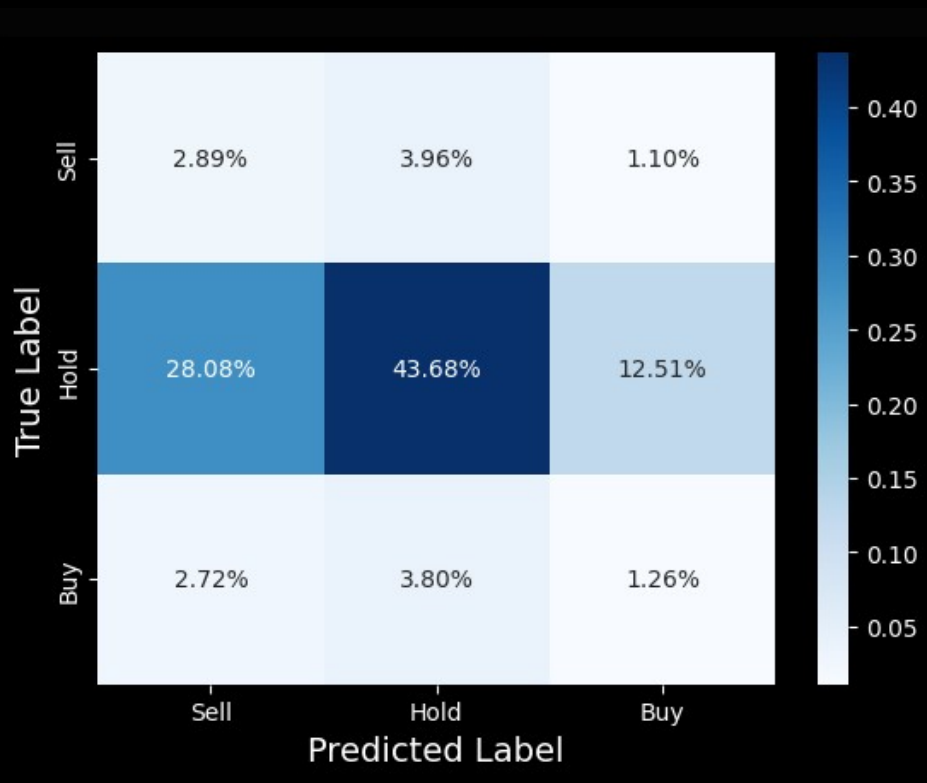


Results: 20years Train 2years Test

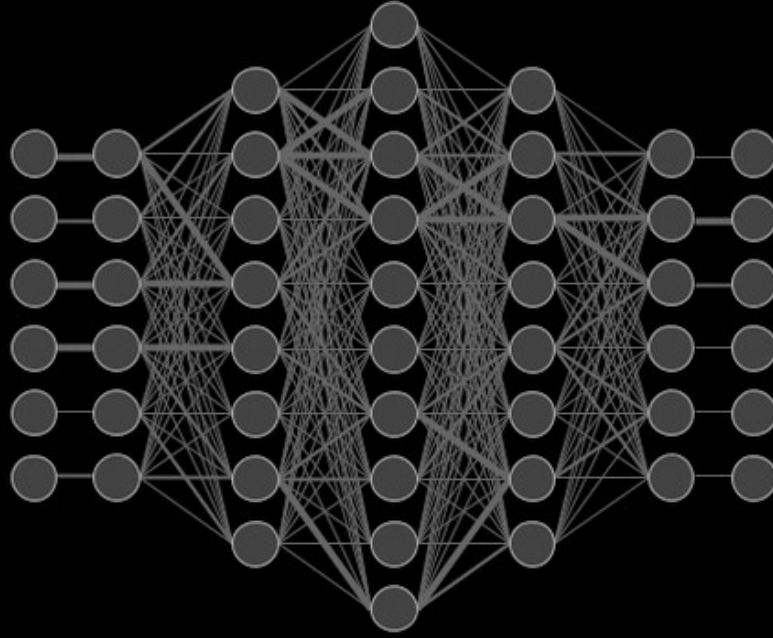
Model accuracy



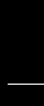
	precision	recall	f1-score	support
Sell	0.0858	0.3634	0.1389	1761
Hold	0.8491	0.5183	0.6437	18647
Buy	0.0848	0.1621	0.1113	1721
accuracy			0.4783	22129
macro avg	0.3399	0.3480	0.2980	22129
weighted avg	0.7290	0.4783	0.5621	22129



CNN better than Logistic regression accuracy score: 0.448



Questions



Appendix: Feature Definitions

BBANDS	Bollinger Bands	ADX	Average Directional Movement Index
DEMA	Double Exponential Moving Average	ADXR	Average Directional Movement Index Rating
EMA	Exponential Moving Average	APO	Absolute Price Oscillator
HT_TRENDLINE	Hilbert Transform - Instantaneous	AROON	Aroon
Trendline		AROONOSC	Aroon Oscillator
KAMA	Kaufman Adaptive Moving Average	BOP	Balance Of Power
MA	Moving average	CCI	Commodity Channel Index
MAMA	MESA Adaptive Moving Average	CMO	Chande Momentum Oscillator
MAVP	Moving average with variable	DX	Directional Movement Index
period		MACD	Moving Average Convergence/Divergence
		MACDEXT	MACD with controllable MA type
AD	Chaikin A/D Line	MACDFIX	Moving Average Convergence/Divergence Fix 12/26
ADOSC	Chaikin A/D Oscillator	MFI	Money Flow Index
OBV	On Balance Volume	MINUS_DI	Minus Directional Indicator
		MINUS_DM	Minus Directional Movement
		MOM	Momentum
		PLUS_DI	Plus Directional Indicator
		PLUS_DM	Plus Directional Movement
ATR	Average True Range	PPO	Percentage Price Oscillator
NATR	Normalized Average True Range	ROC	Rate of change : ((price/prevPrice)-1)*100
TRANGE	True Range	ROCP	Rate of change Percentage: (price-prevPrice)/prevPrice
		ROCR	Rate of change ratio: (price/prevPrice)
		ROCR100	Rate of change ratio 100 scale: (price/prevPrice)*100
		RSI	Relative Strength Index
BETA	Beta	STOCH	Stochastic
CORREL	Pearson's Correlation	STOCHF	Stochastic Fast
Coefficient (r)		STOCHRSI	Stochastic Relative Strength Index
LINEARREG	Linear Regression	TRIX	1-day Rate-Of-Change (ROC) of a Triple Smooth EMA
LINEARREG_ANGLE	Linear Regression Angle	ULTOSC	Ultimate Oscillator
LINEARREG_INTERCEPT	Linear Regression Intercept	WILLR	Williams' %R
LINEARREG_SLOPE	Linear Regression Slope		
STDDEV	Standard Deviation		
TSF	Time Series Forecast		
VAR	Variance		