

Team 42 Final Project - LSTM and CNN Based Stock Market Predictor

Kyu, AM; Kyu, AW; McKnight, MP

ABSTRACT

Predicting the stock market is a challenging task that has been attempted with a variety of methods. In this project, a neural network model that combines aspects of both CNNs and LSTMs to predict the appropriate position of a company's stock (Buy, Sell, or Hold) using stock data and technical indicators commonly used by analysts. The proposed model outperformed the baseline model, but still yielded unfavorable results.

THE DATA

The data was collected from Yahoo Finance’s historical stock data, which includes open price, high, low, closing price, adjusted closing price, and volume. The data was further augmented by creating over 150 technical indicators using the TA-Lib Library. The indicators fall within the following groups: overlap studies, momentum indicators, volume indicators, volatility indicators, price transform, cycle indicators, and pattern recognition. The indicators were separated depending on if they were periodic or non-periodic. The periodic indicators were calculated over a range of periods from 5 to 20. The periodic data was then arranged into images [4], where the row is the period and the columns are indicators to be inputted in the CNN AE. Due to the overall upward trend for AMD the hold label dominated the classes. To combat this we downsampled this dominant output to have a more general model that could translate to stocks that are less consistently growing or are in a downward trend.

Pattern Recognition		Momentum Indicators	
COL2CROWS	Two Crows	ADX	Average Directional Movement Index
COL3BLACKCROWS	Three Black Crows	ADXR	Average Directional Movement Index Rating
COL3INSIDE	Three Inside up/Down	APO	Absolute Price Oscillator
COL3LINESTRIKE	Three-Line Strike	AROON	Aroon
COL3OUTSIDE	Three Outside up/Down	AROONOSC	Aroon Oscillator
COL3STARSINSOUTH	Three Stars In The South	BOP	Balance of Power
COL3WHITESOLDIERS	Three Advancing white soldiers	CCI	Commodity channel index
COLABANDONEDBABY	Abandoned Baby	CNO	Chande Momentum Oscillator
COLADVANCEBLOCK	Advance Block	DX	Directional Movement Index
COLBELTHOLD	Belt-hold	HACD	Moving Average Convergence/Divergence
COLBREAKAWAY	Breakaway	HACDEXT	HACD with controllable MA type
COLCLOSINGHARUBOZU	Closing Harubozu	HACDFIX	Moving Average Convergence/Divergence Fix 12/26
COLCONCEALBABYSWALL	Concealing Baby Swallow	MFI	Money Flow Index
COLCOUNTERATTACK	Counterattack	MINUS_DI	Minus Directional Indicator
COLDARKCLOUDCOVER	Dark Cloud Cover	MINUS_DM	Minus Directional Movement
COLDOJI	Doji	MOI	Momentum
COLDOJISTAR	Doji Star	PLUS_DI	Plus Directional Indicator
COLDRAGONFLYDOJI	Dragonfly Doji	PLUS_DM	Plus Directional Movement
COLENGULFING	Engulfing Pattern	PPO	Percentage Price Oscillator
COLERVENINGDOJISTAR	Evening Doji Star	ROC	Rate of change : ((price/prevPrice)-1)*100
COLERVENINGSTAR	Evening Star	ROCP	Rate of change Percentage: (price-prevPrice)/prevPrice
COLGAPSIDESIDEWHITE	Up/Down-gap side-by-side white lines	ROCR	Rate of change ratio: (price/prevPrice)
COLGRAVESTONEDOJI	Gravestone Doji	ROCR100	Rate of change ratio 100 scale: (price/prevPrice)*100
COLHAMMER	Hammer	RSI	Relative Strength Index
COLHANGINGMAN	Hanging Man	STOCH	Stochastic
COLHARAMI	Harami Pattern	STOCHF	Stochastic Fast
COLHARAMICROSS	Harami Cross Pattern	STOCHRSI	Stochastic Relative Strength Index
COLHIGMAVE	High-wave Candle	TRIX	1-day Rate-of-Change (ROC) of a Triple Smooth EMA
COLHIKKAKE	Hikikake Pattern	ULTOSC	Ultimate Oscillator
COLHIKKAKEPICK	Modified hikikake Pattern	WILLR	Williams %R
COLHOMINGPIGION	Homing Pigeon	Overlap Studies	
COLIDENTICAL3CROWS	Identical Three Crows	BBANDS	Bollinger Bands
COLINNECK	In-Neck Pattern	DEMA	Double Exponential Moving Average
COLINVERTEDHAMMER	Inverted Hammer	EMA	Exponential Moving Average
COLKICKING	Kicking	HT_TRENDLINE	Hilbert Transform - Instantaneous Trendline
COLKICKINGBYLENGTH	Kicking - bull/bear determined by the longer marubozu	KAMA	Kaufman Adaptive Moving Average
COLLADDERBOTTOM	Ladder Bottom	MA	Moving average
COLLONGLEGGEDDOJI	Long Legged Doji	MAMA	MESA Adaptive Moving Average
COLLONGLINE	Long Line Candle	MAVP	Moving average with variable period
COLMARUBOZU	Marubozu	MIDPOINT	Midpoint over period
COLMATCHINGLOW	Matching Low	MIDPRICE	Midpoint Price over period
COLMATHOLD	Mat hold	SAR	Parabolic SAR
COLMORNINGDOJISTAR	Morning Doji Star	SAREXT	Parabolic SAR - Extended
COLMORNINGSTAR	Morning Star	SPA	Simple Moving Average
COLNECK	On-Neck Pattern	T3	Triple Exponential Moving Average (T3)
COLPIERCING	Piercing Pattern	TEMA	Triple Exponential Moving Average
COLRICKSHAWMAN	Rickshaw Man	TRIMA	Triangular Moving Average
COLRISEFALLMETHODS	Rising/Falling Three Methods	WMA	Weighted Moving Average
COLSEPARATINGLINES	Separating Lines	Volume Indicators	
COLSHOOTINGSTAR	Shooting Star	AO	Chaikin A/D Line
COLSHORTLINE	Short Line Candle	AOSC	Chaikin A/D Oscillator
COLSPINNINGTOP	Spinning Top	OBV	On Balance Volume
COLSTALLEDPATTERN	Stalled Pattern	Cycle Indicators	
COLSTICKSANDWICH	Stick Sandwich	HT_DPERIOD	Hilbert Transform - Dominant Cycle Period
COLTAURIT	Taukuri (Dragonfly Doji with very long lower shadow)	HT_DPHASE	Hilbert Transform - Dominant Cycle Phase
COLTASUKIGAP	Tasuki Gap	HT_PHASOR	Hilbert Transform - Phasor Components
COLTHRUSTING	Thrusting Pattern	HT_SINE	Hilbert Transform - Sinewave
COLTRISTAR	Tristar Pattern	HT_TRENDMODE	Hilbert Transform - Trend vs Cycle Mode
COLUNIQUE3RIVER	Unique 3 River	Price Transform	
COLUPSIDEGAP2CROWS	Upside Gap Two Crows	AVERAGE	Average Price
COLXSIDEGAP3METHODS	Upside/Downside Gap Three Methods	MEDPRICE	Median Price
Statistic Functions		TYPPRICE	Typical Price
BETA	Beta	WCLPRICE	Weighted Close Price
CORREL	Pearson's Correlation Coefficient (r)	Volatility Indicators	
LINEARREG	Linear Regression	ATR	Average True Range
LINEARREG_ANGLE	Linear Regression Angle	NATR	Normalized Average True Range
LINEARREG_INTERCEPT	Linear Regression Intercept	TRANGE	True Range
LINEARREG_SLOPE	Linear Regression Slope	Data From Yahoo Finance	
STDEV	Standard Deviation	open	open
TSP	Time Series Forecast	close	close
VAR	Variance	high	high
		low	low
		adjusted close	adjusted close
		volume	volume

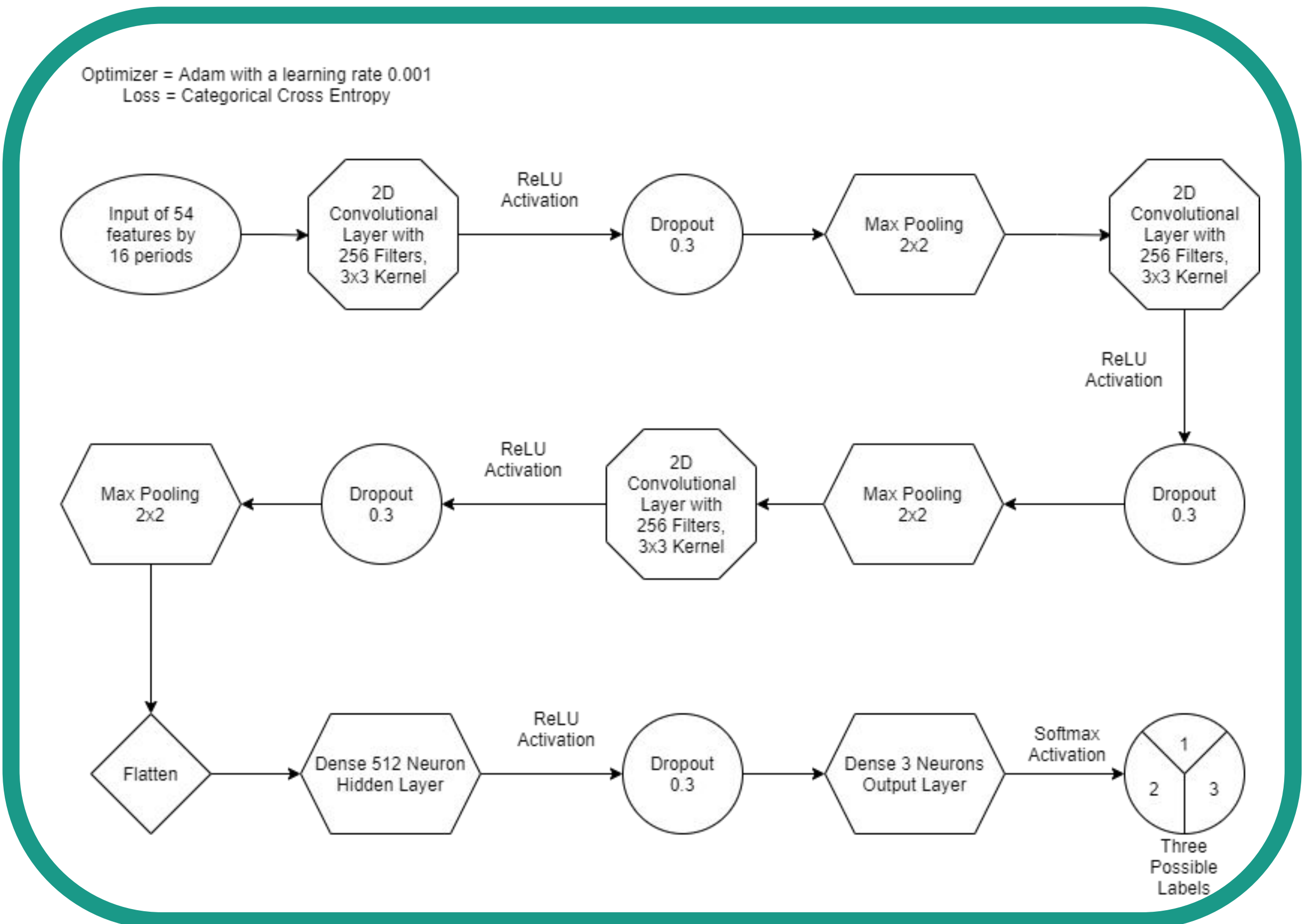
TASK & MOTIVATION

Motivation: Predicting the stock market has always been a classical neural network problem. However, very few, if any, models (CNN or LSTM) have been very successful in predicting outcomes of the stock market accurately [2]. Therefore, we propose a new neural network model based off of literature that combines a convolutional neural network with a LSTM neural network model in hopes to surpass the results of current models [3].

Task: To predict appropriate buy, sell, or hold positions on the price of a specific company’s stock (Advanced Micro Devices - AMD) in order to improve returns from day trading.

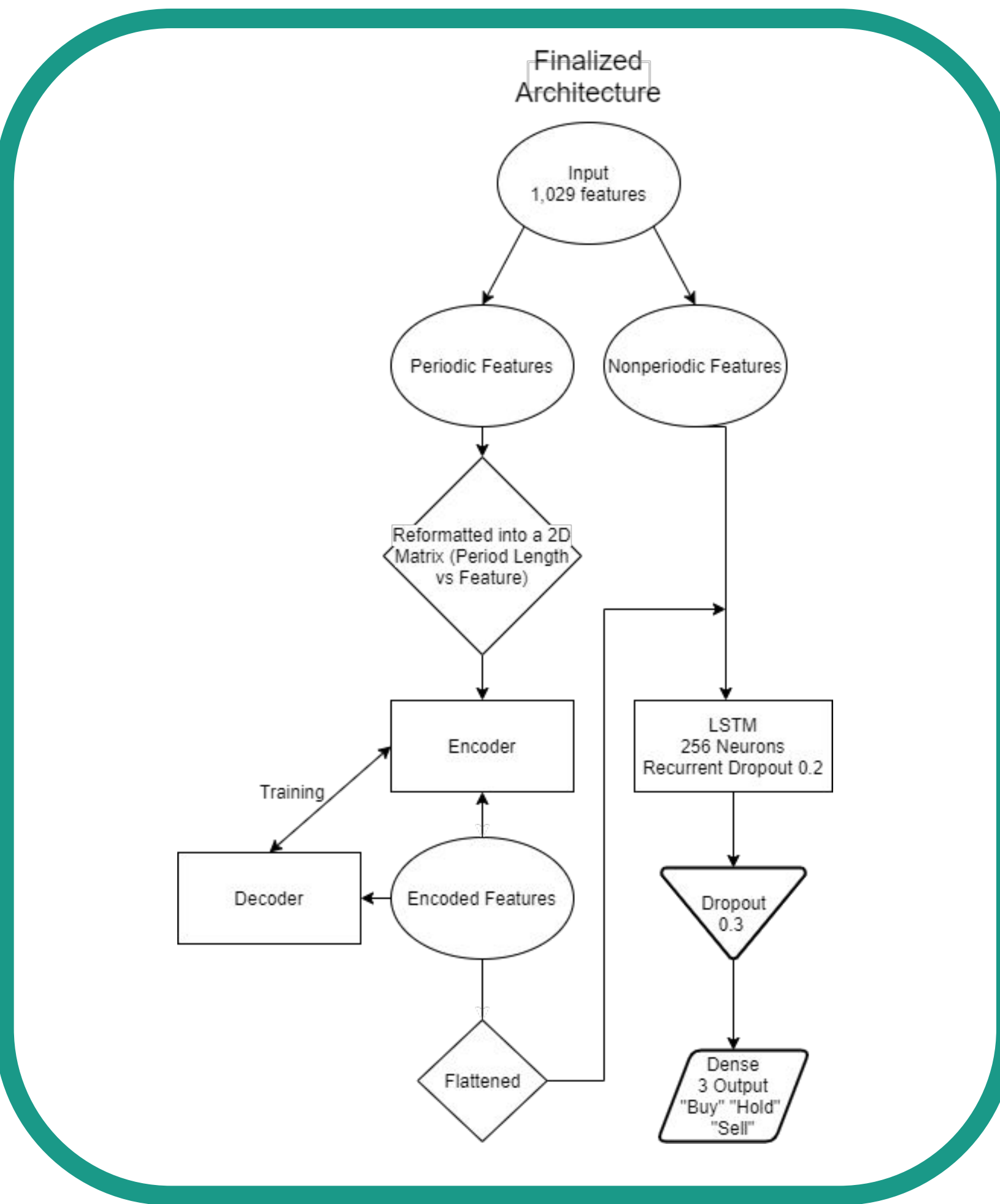
MODEL ARCHITECTURE

The following models were optimized using the Talos library.
Baseline: Deep CNN Model with Dropout and Max Pooling Layers



Proposed Model:

The goal of this architecture was to not only take advantage of non-periodic data, as in the baseline, but to also use periodic features. This is done by having an autoencoder network reduce the dimensionality of the periodic features so it can be combined with the non-periodic features and then fed to the LSTM network for classification.



RESULTS

Computational Model Evaluation
Baseline Model:

Class	Precision	Recall	F1-Score	Accuracy
Sell (0)	0.310	0.562	0.400	0.5625
Buy (1)	0.238	0.714	0.357	0.7143
Hold (2)	0.943	0.764	0.844	0.7639
Total Accuracy				0.748

Proposed Model:

Class	Precision	Recall	F1-Score	Accuracy
Sell (0)	0	0	0	0
Buy (1)	0	0	0	0
Hold (2)	0.878	1.000	0.935	1.00
Total Accuracy				0.878

Financial Evaluation

For the financial evaluation, we compared our models performance in the past year to two other values: annualized percent gain from long term trading and maximum percent gain.

	Baseline Model	Proposed Model	Long-term Trading	Maximum Percent Gain
Total Percent Gain (%)	-16.36%	0%	50.65%	133.11%

NEXT STEPS

- Expand to other stock symbols or create more general model
- Implement NLP for stock related news analysis
- Explore various filtering methods for the data
- Investigate interdependencies of companies in similar sectors

REFERENCES

[1] NASDAQ. (2021, April 25). Advanced Micro Devices, Inc. (AMD). [Stock quote]. Retrieved from <https://finance.yahoo.com/quote/AMD>

[2] D. Nelson, A. M. Pereira, and Renato, "Stock market's price movement prediction with LSTM neural networks", 2017 International Joint Conference on Neural Networks (IJCNN), 2017. doi: 10.1109/IJCNN.2017.7966019 (accessed Feb. 27, 2021).

[3] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory | Neural Computation," Neural Computation, 2020. <https://doi.org/10.1162/neco.1997.9.8.1735> (accessed Feb. 27, 2021).

[4] O. B. Sezer and M. Ozbayoglu, "Algorithmic Financial Trading with Deep Convolutional Neural Networks: Time Series to Image Conversion Approach," Applied Soft Computing, vol. 70, pp. 525–538, Sep. 2018. <https://doi.org/10.1016/j.asoc.2018.04.024>