Ensemble Deep Learning Framework for Financial Time Series

Saurabh Mishra, PhD

with Professor Bilal Ayyub

University of Maryland 1856
A. James Clark School of Engineering
Department of Civil & Environmental Engineering

CTSM
Center for Technology and Systems Management
Technology for intelligent Decisions

Outline

- Problem Statement
- Methods
- Research Objectives
- Proposed Framework
- Results
- Forecast Monitoring
- Conclusions

Monitoring Opportunities and Threats in Real-Time

- **Opportunities and Threats –** from geo-political to economic, finance to climate
- **Risk perspectives –** multiple stakeholders from different departments are involved
- **Mass movements –** people are interested to monitor large shifts in prices, behavioral fluctuations, social and economic trends
- **Monitoring Technologies –** AI is transforming how business is done but little application to complex dynamic systems and policy-making has not yet been deployed
- **Executive Decision-Making –** Executives need applications to help them make informed-decisions about their investments and policy moves

Data-driven identification of Opportunities and Threats

Cross Sector Risk Domains and Impacts

Geopolitics
opportunities: defense negotiations, trade agreements
threats: terror attacks, revolutions, protests

Economics
opportunities: fast economic growth, decrease in public debt, job growth
threats: financial crisis, economic contraction, rising unemployment

Finance
opportunities: stock index or ticker growth, stable EXR, bitcoin price growth
threats: stock market crash, exchange rate instability, cryptocurrency crash

Methods

- Clustering
  Dynamic Time Warping, Hierarchical Clustering, K-means, t-SNE, MDS, UMAP

- Converting Time-Series to Class Labels
  Z-threshold peak detection algorithm, Bayesian change-point detection,

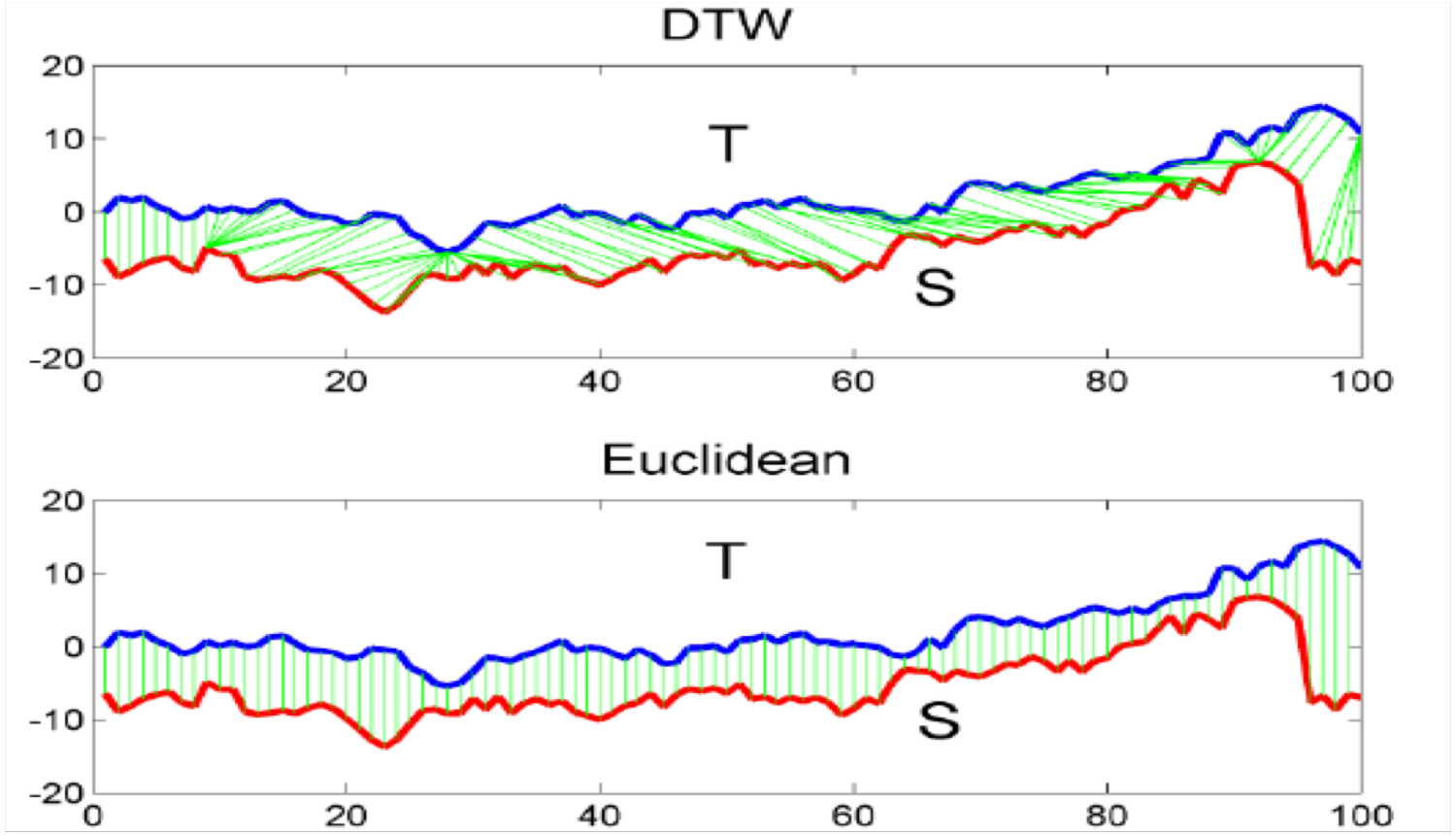
- Classification
  Naïve Bayes Classifier, Support Vector Machines (SVMs), Convolutional Neural Networks (CNNs)

- Time-Series Sequence Prediction
  ARMA, ARIMA, VAR, Kalman Filter, Particle Filter, Recurrent Neural Networks (RNNs), Long-Short Term Memory (LSTM)

- Uncertainty Quantification (UQ)
  MVE, Bootstrap, Bayesian, MC Dropout, MQ-RNN, Tilt Loss Function

Difference between Dynamic Time Warping (DTW) distance and Euclidean distance

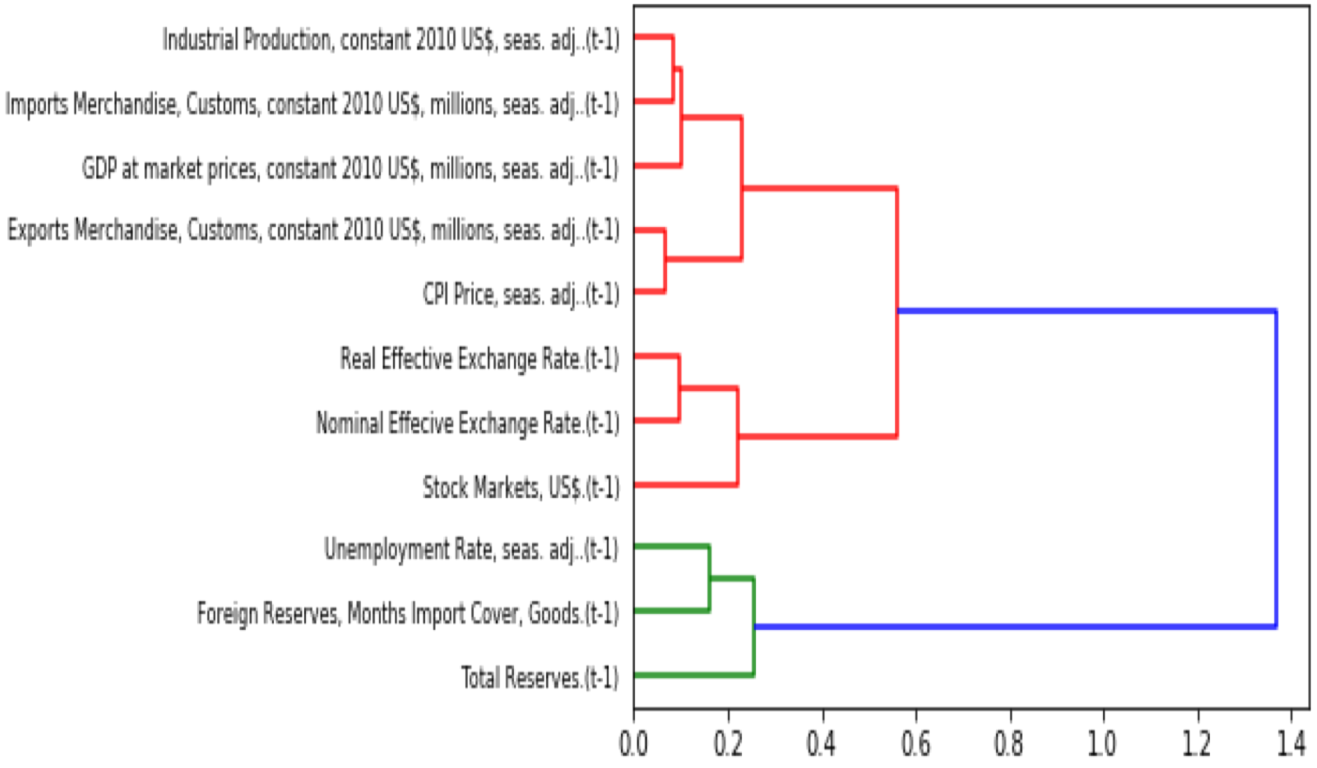
(green lines represent mapping between points of time series T and S).

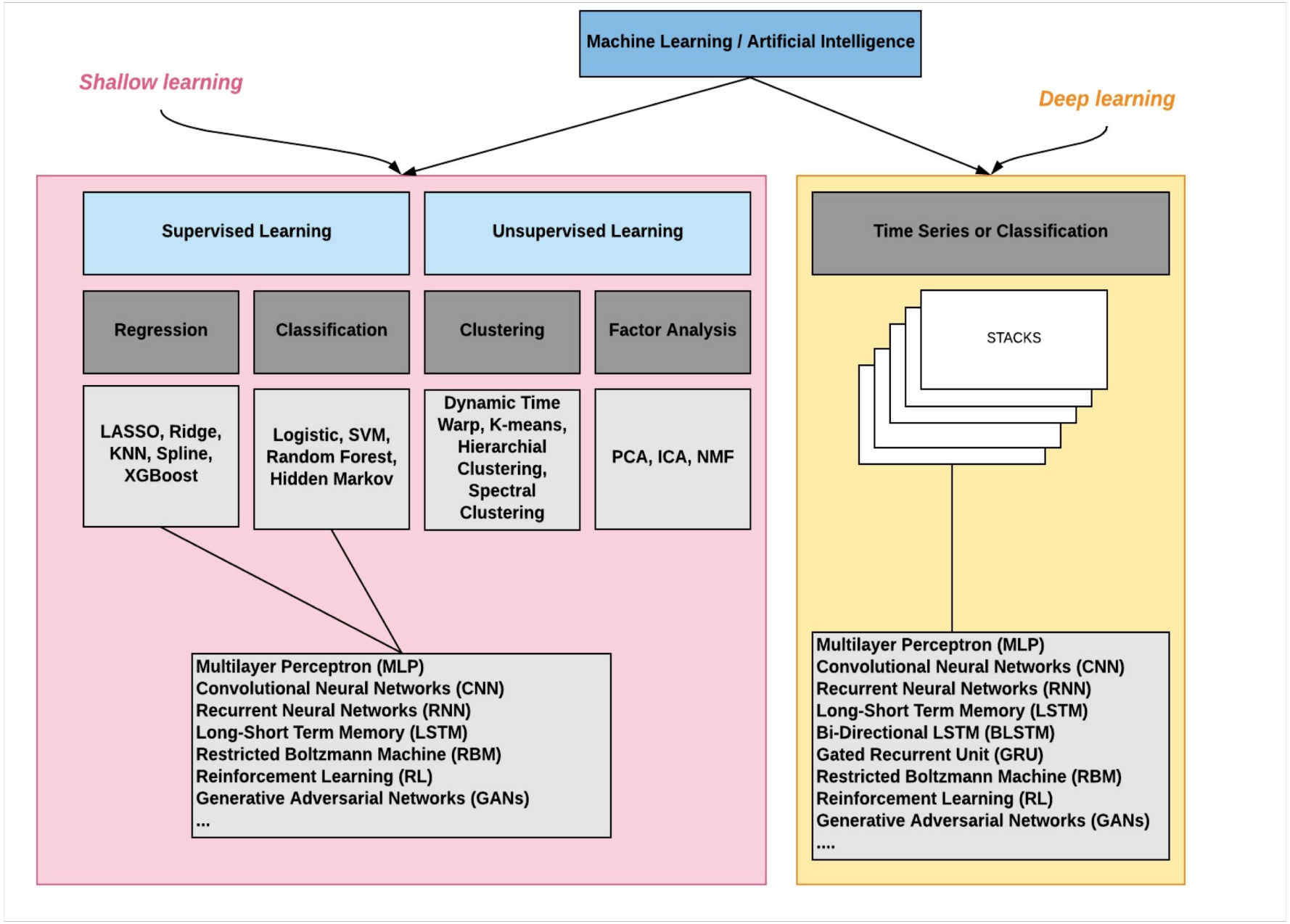


Source: Cassisi et al. 2012.

Hierarchical Clustering Analysis (HCA)

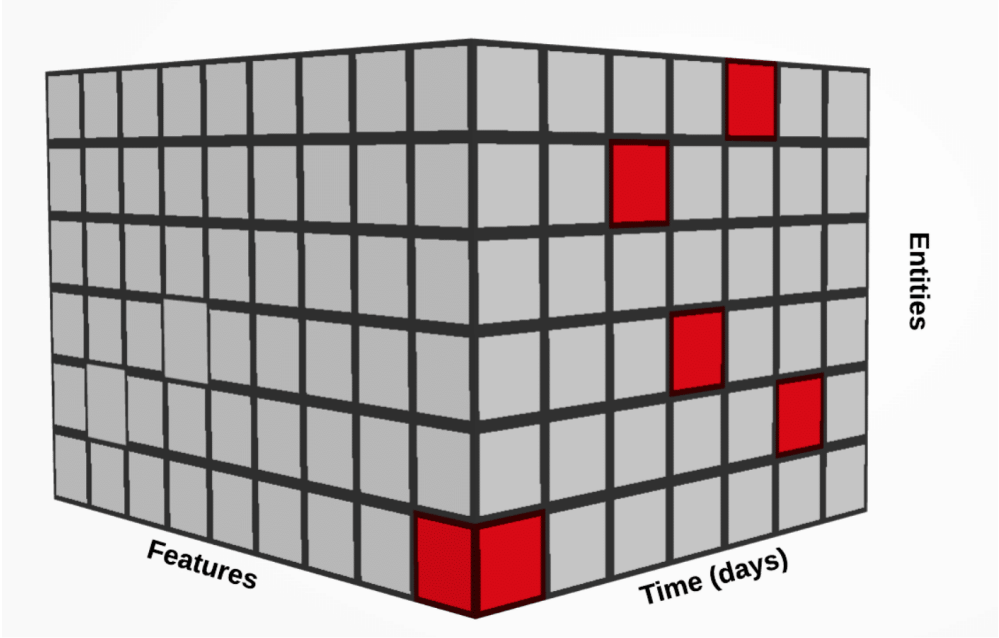
HCA example for
macroeconomic variables, USA



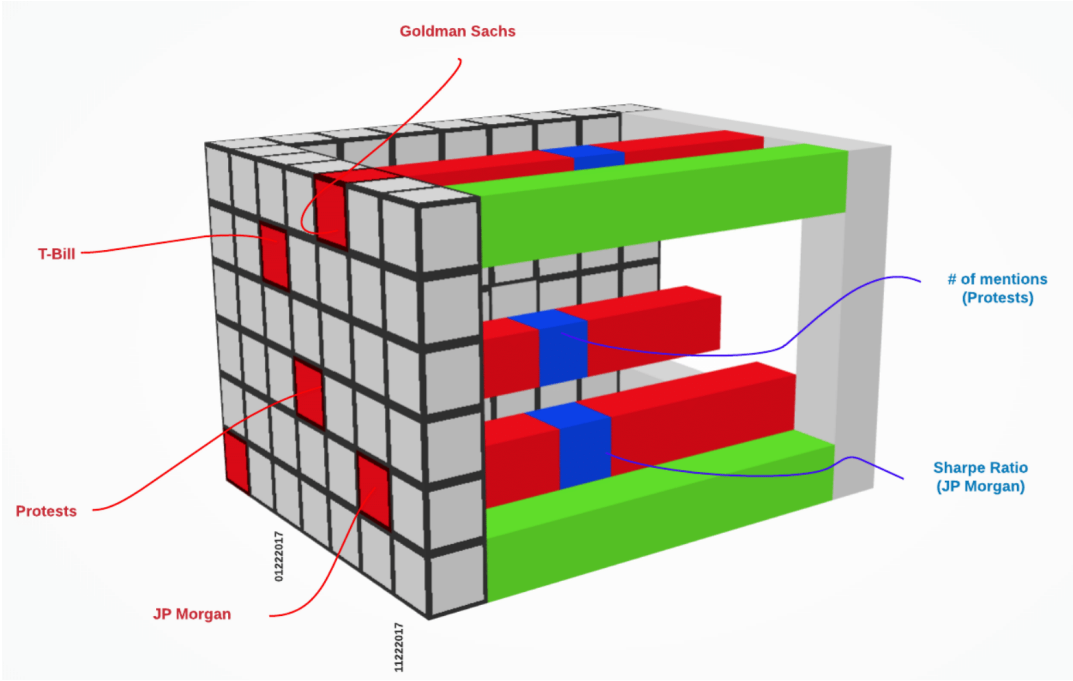

Machine Learning / Artificial Intelligence
Shallow learning
Deep learning
Supervised Learning
Unsupervised Learning
Time Series or Classification
Regression
Classification
Clustering
Factor Analysis
LASSO, Ridge, KNN, Spline, XGBoost
Logistic, SVM, Random Forest, Hidden Markov
Dynamic Time Warp, K-means, Hierarchial Clustering, Spectral Clustering
PCA, ICA, NMF
STACKS
Multilayer Perceptron (MLP)
Convolutional Neural Networks (CNN)
Recurrent Neural Networks (RNN)
Long-Short Term Memory (LSTM)
Restricted Boltzmann Machine (RBM)
Reinforcement Learning (RL)
Generative Adversarial Networks (GANs)
...
Multilayer Perceptron (MLP)
Convolutional Neural Networks (CNN)
Recurrent Neural Networks (RNN)
Long-Short Term Memory (LSTM)
Bi-Directional LSTM (BLSTM)
Gated Recurrent Unit (GRU)
Restricted Boltzmann Machine (RBM)
Reinforcement Learning (RL)
Generative Adversarial Networks (GANs)
....

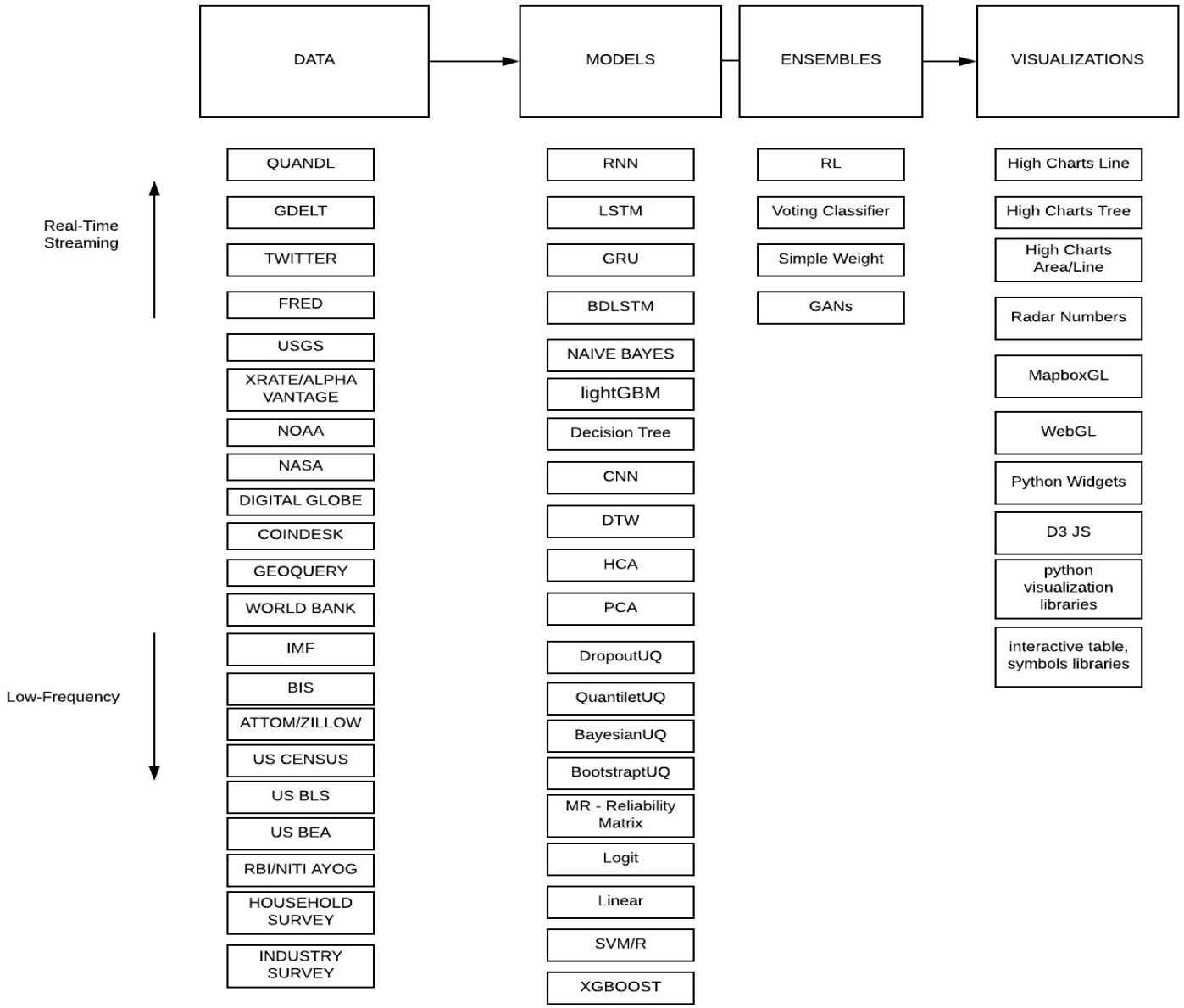
Database

3d Database

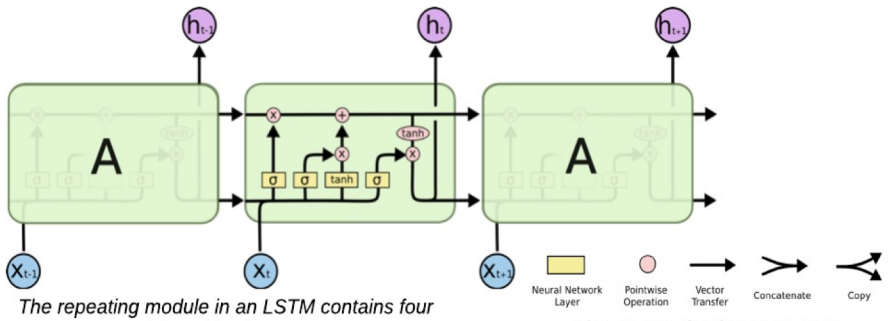


Abstraction of entity, time, and features




DATA
MODELS
ENSEMBLES
VISUALIZATIONS
Real-Time Streaming
Low-Frequency
QUANDL
GDELT
TWITTER
FRED
USGS
XRATE/ALPHA VANTAGE
NOAA
NASA
DIGITAL GLOBE
COINDESK
GEOQUERY
WORLD BANK
IMF
BIS
ATTOM/ZILLOW
US CENSUS
US BLS
US BEA
RBI/NITI AYOG
HOUSEHOLD SURVEY
INDUSTRY SURVEY
RNN
LSTM
GRU
BDLSTM
NAIVE BAYES
lightGBM
Decision Tree
CNN
DTW
HCA
PCA
DropoutUQ
QuantiletUQ
BayesianUQ
BootstraptUQ
MR - Reliability Matrix
Logit
Linear
SVM/R
XGBOOST
RL
Voting Classifier
Simple Weight
GANs
High Charts Line
High Charts Tree
High Charts Area/Line
Radar Numbers
MapboxGL
WebGL
Python Widgets
D3 JS
python visualization libraries
interactive table, symbols libraries

Long-Short Term Memory Neural Network

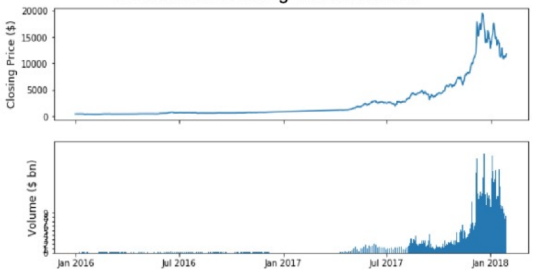


The repeating module in an LSTM contains four interacting layers.
Source: Cohl Blog 2015.

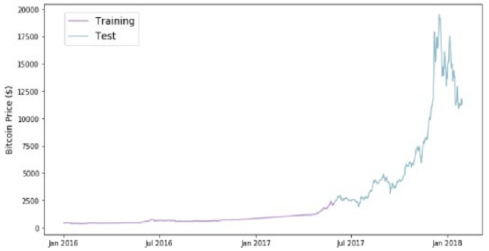
Sign Denotation in LSTM RNN

Simple example of LSTM in bitcoin price prediction

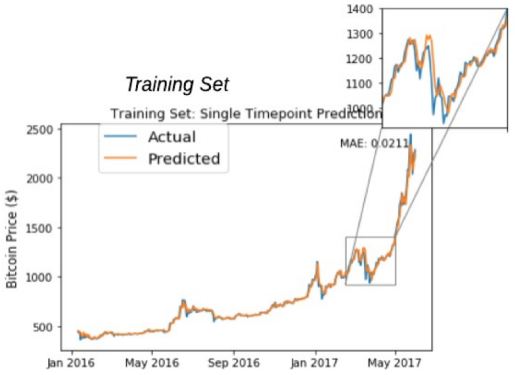
Raw Bitcoin Closing Dat Price data



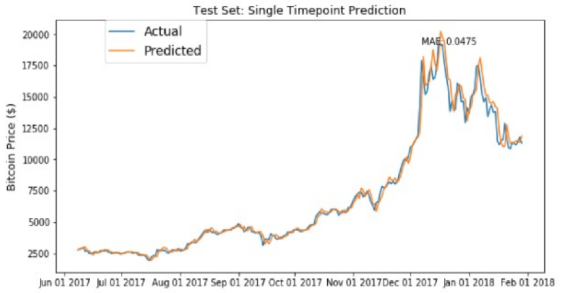
Split in Training and Test Set



Training Set



Test Set



Research Objectives

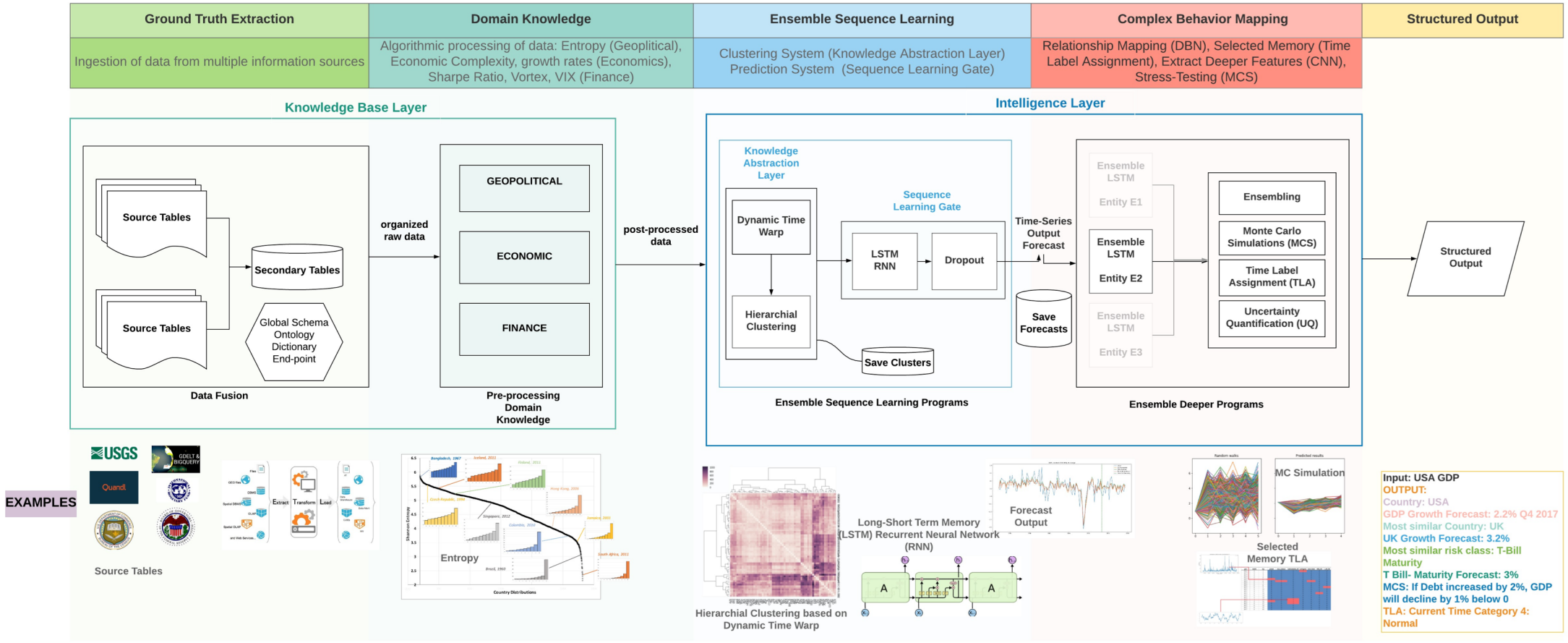
- Structure Domain Logic for Training Data: Use unsupervised machine learning models to build training data

- Establish the proposed model framework in context of multi-variate time-series sequence prediction problem for domains of interest

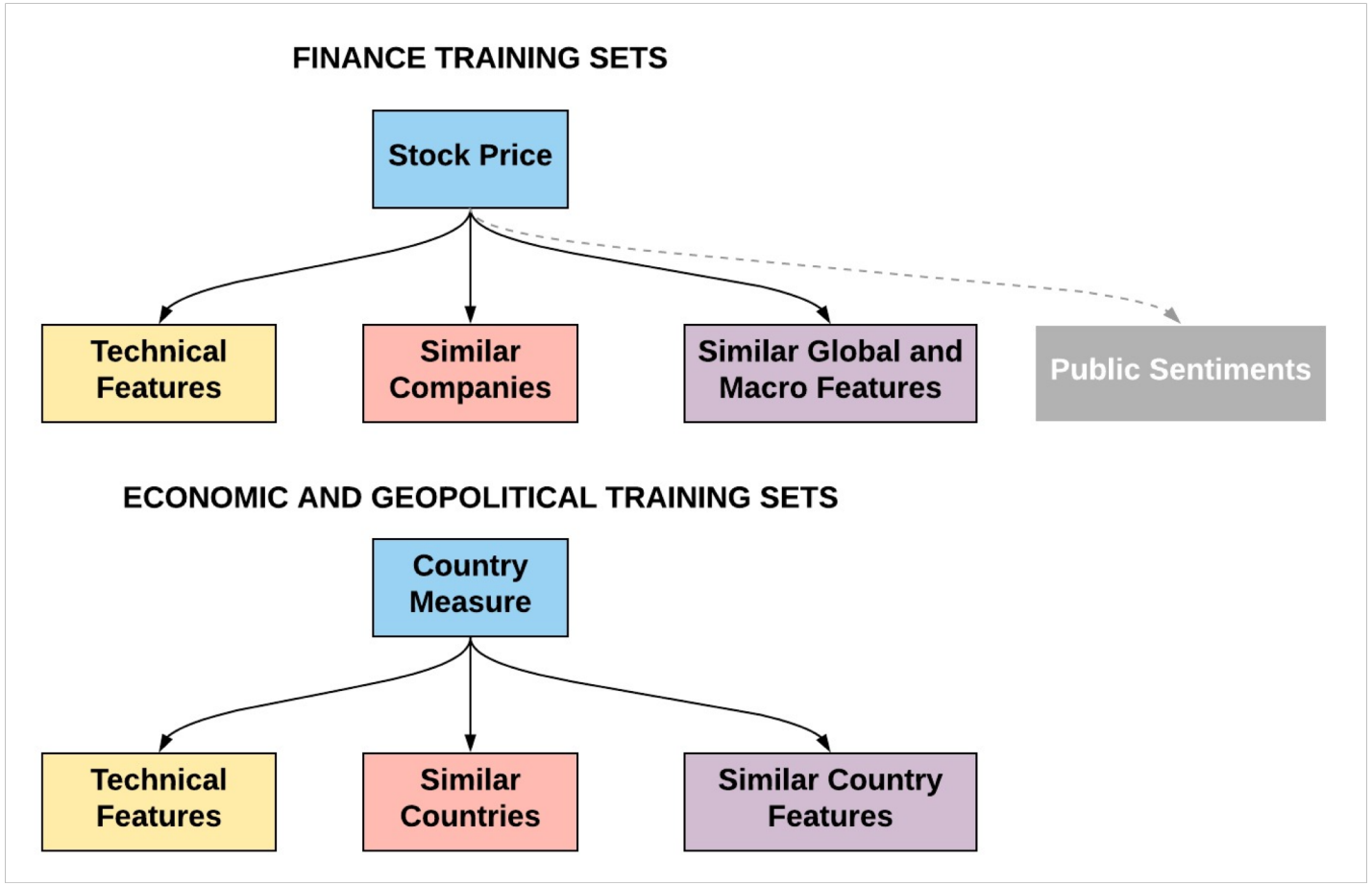
- Establish in which cases proposed model framework works better than others

- Try to go beyond Black Box and provide examples to assess viability of proposed model framework through:
  - Uncertainty Quantification (UQ)
  - Benchmark Assess Model Reliability (MR)
  - Simulations/External Shocks
  - Alert System

Proposed Model Framework



Preprocessing Domain Knowledge and Structured Logic Approach to Training Groups

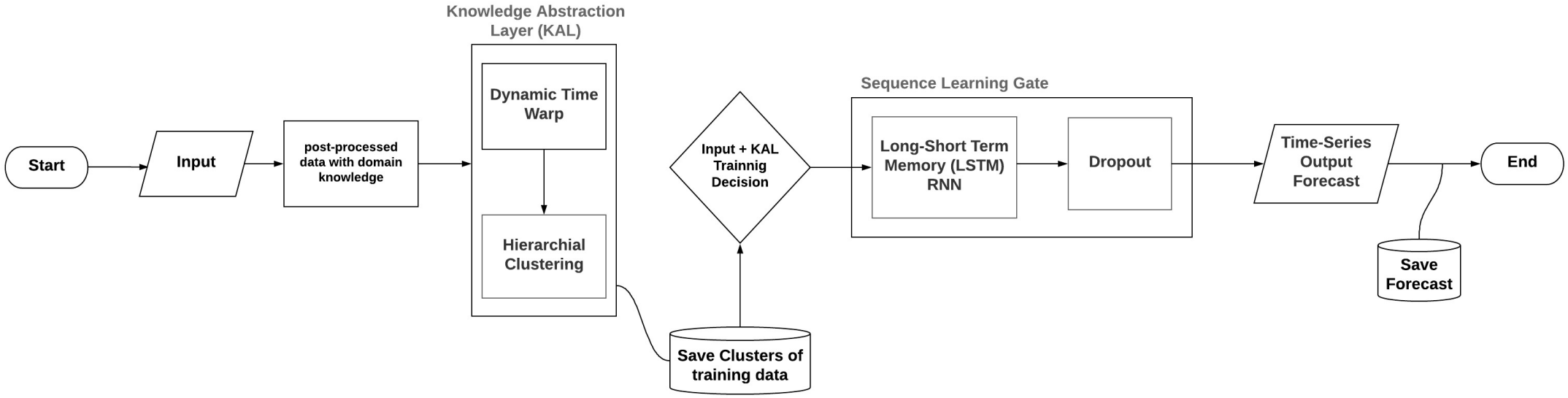


Processing Domain Knowledge

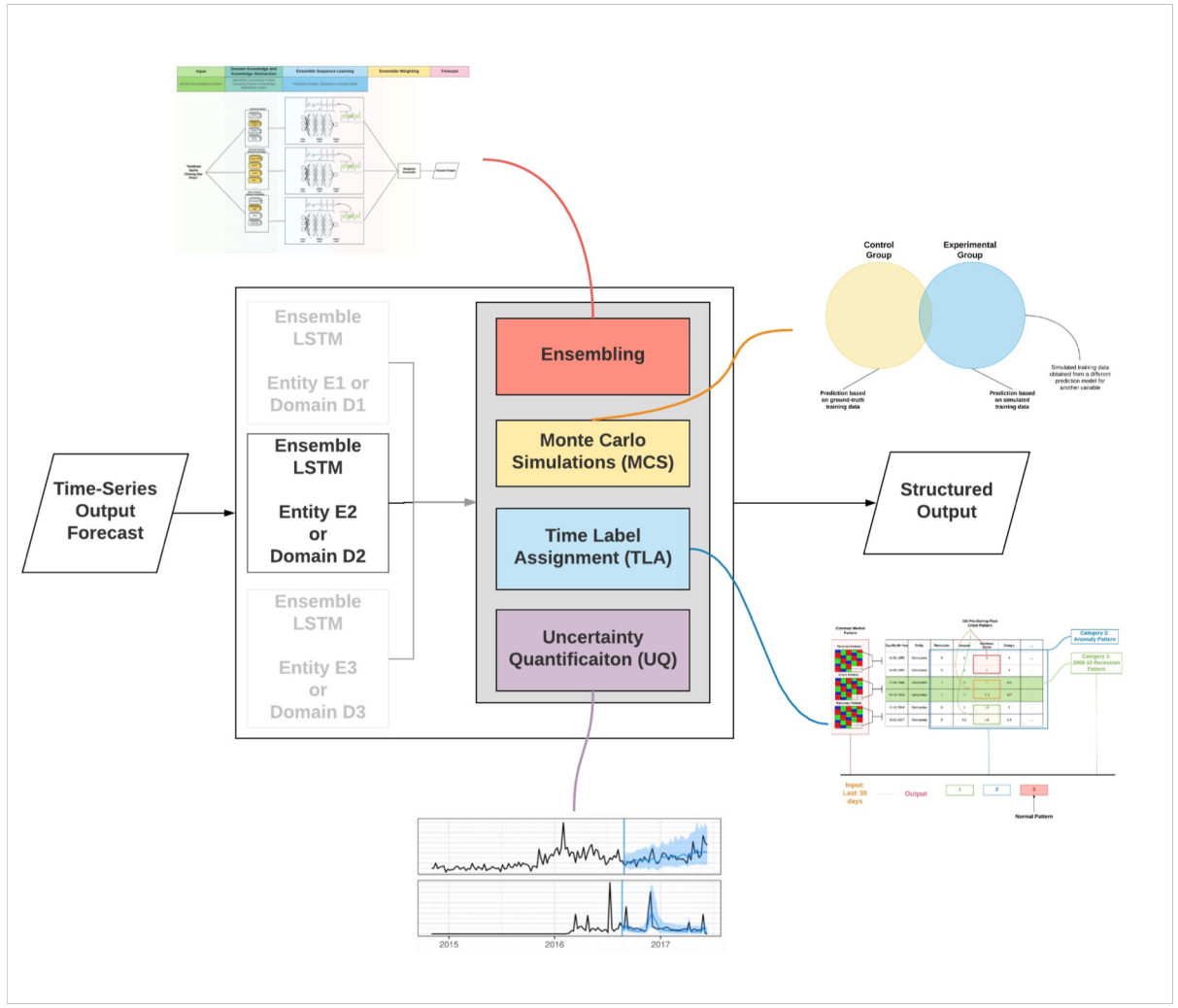
Take given input data stream, using domain-specific methods can be extremely valuable to extract many new features, for example

- Macroeconomics: Economic Diversification Measures, Production Sophistication Measures, Income Inequality, etc.
- Geopolitics: Tone, Volume, Goldstein Score, etc.
- Finance: Technical Features (30+), Similar Indices, Economic Foundations, Sentiments

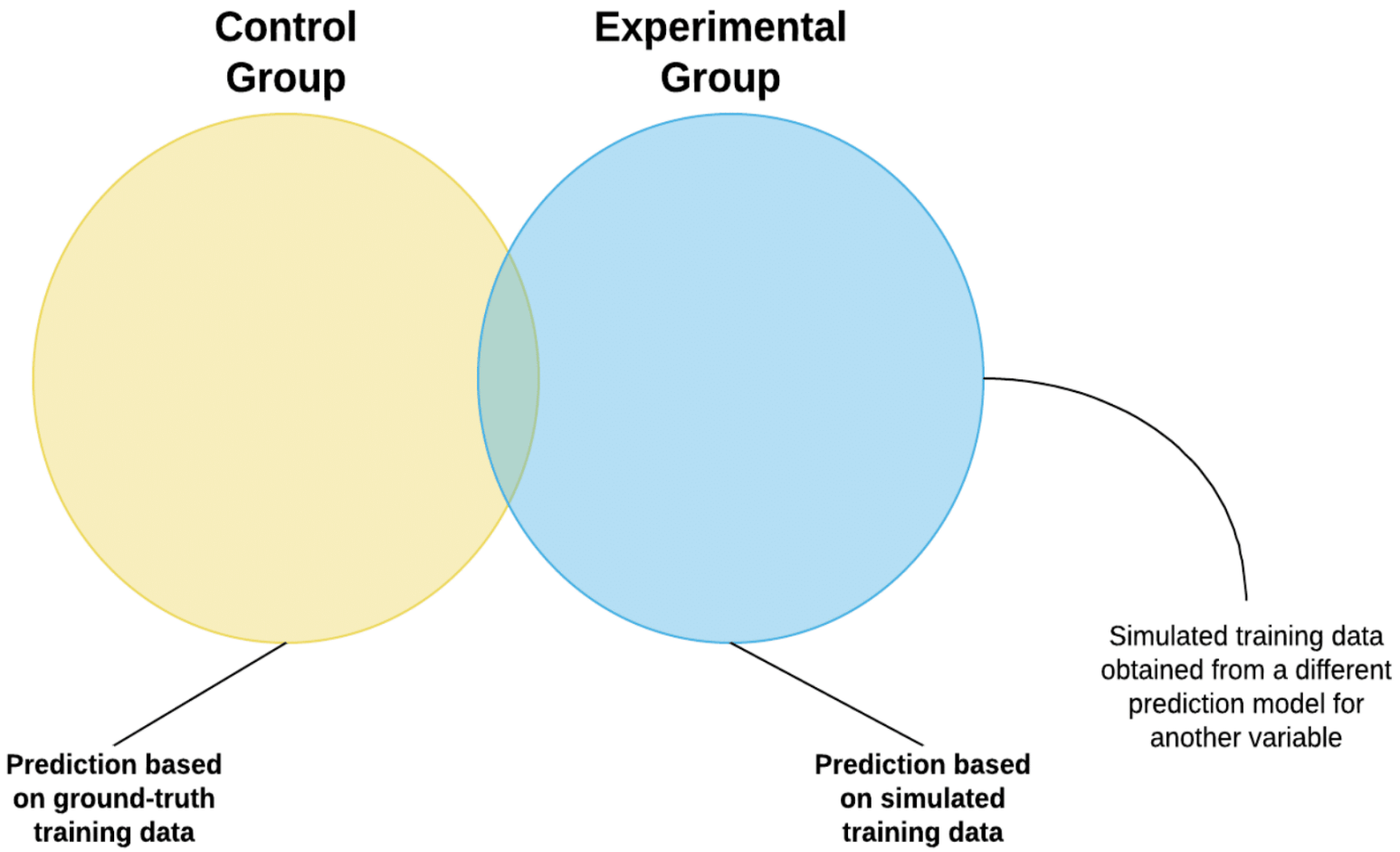
Sequence Learning Prediction Framework



Go Beyond Black Box to Provide Structured Output



Monte Carlo Simulation (MCS) for a more controlled-environment for stress-testing

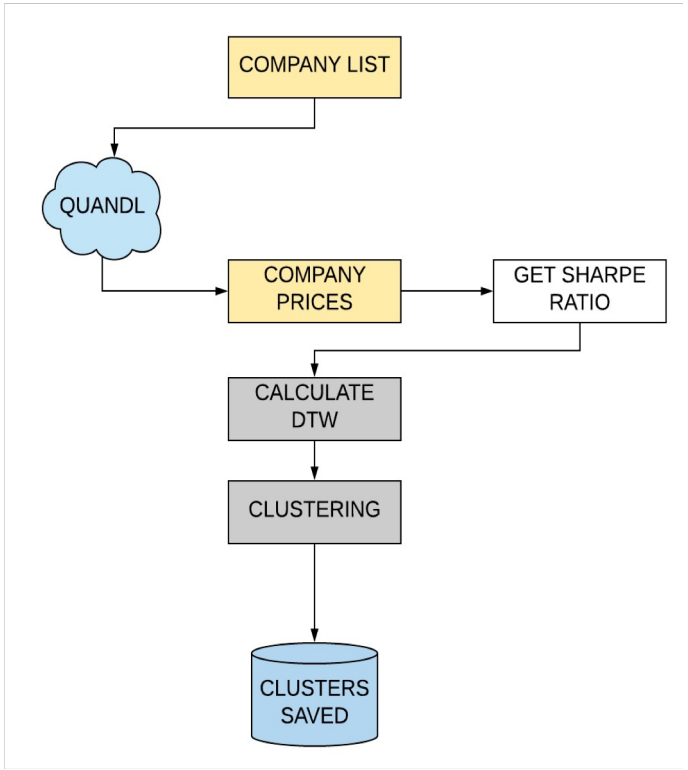


Uncertainty Quantification

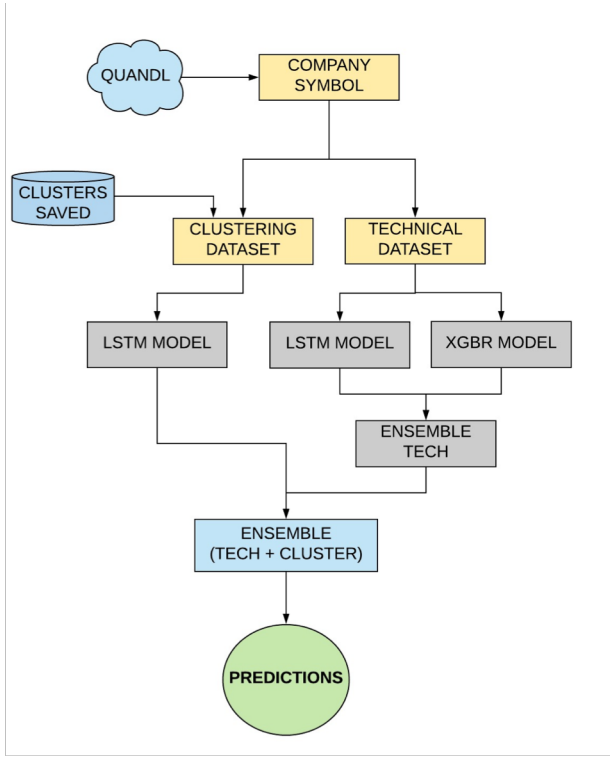
- Mean Variance Estimation
- Bootstrap
- Bayesian Methods
- **MC Dropout method**
- MQ-RNN
- Pinball Loss Function

Papers	Methods	Summary
Gal, Y. and Z. Ghahramani, 2015; Gal, Y. and Z. Ghahramani, 2016a; Gal, Y. and Z. Ghahramani, 2016b; Gal, Y, 2016	Combining modern deep learning and Bayesian techniques for uncertainty quantification, variational inference, MC Dropout as a Bayesian approximator	Bayesian method is computationally expensive, but MC dropout framework can estimated for almost free to obtain prediction intervals
Wen, Torkkola, Narayanaswamy, 2017	Multi-horizon Quantile Recurrent Forecaster (MQ-RNN), large-scale time-series forecasting at Amazon	Nonparametric nature of Quantile Regression and the efficiency of Direct Multi-Horizon Forecasting is scalable but computationally expensive in network framework
Khosravi et al. 2014	Delta, Classical, MVE, Bootstrap, Bayesian methods comparisons, Genetic Algorithm for ensembling various UQ methods	Comprehensive review and comparison, ensemble method is promising and important for system architecture UQ modularity
Solimum 2014	Empirical comparison of Classical, Bayesian and Bootstrap methods for UQ	Bootstrap method has the smallest width confidence interval, thereby more precise than Bayesian and Classical methods
Zhu, L., and N. Laptev, 2017	Deep and Confident Prediction for Time Series at Uber, real-time anomaly detection across millions of metrics, Bayesian deep LSTM for time-series prediction and uncertainty	Computationally expensive and useful for anomaly detection, method is essentially the same as Gal 2015, Gal and Ghahramani 2016, however, a computationally intensive encoder-decoder neural network framework
Abeywardana and Ramos 2015	Quantile Regression Loss Function application MLP, our application applies to RNN and LSTMs	Non-parametric method of inferring quantiles and derivation a novel Variational Bayesian (VB) approximation to the marginal likelihood, similar to Gal 2015, Gal and Ghahramani 2016, however using a tilt (pinball) loss function, easier implementation

Clustering and Prediction System for Stock Price data

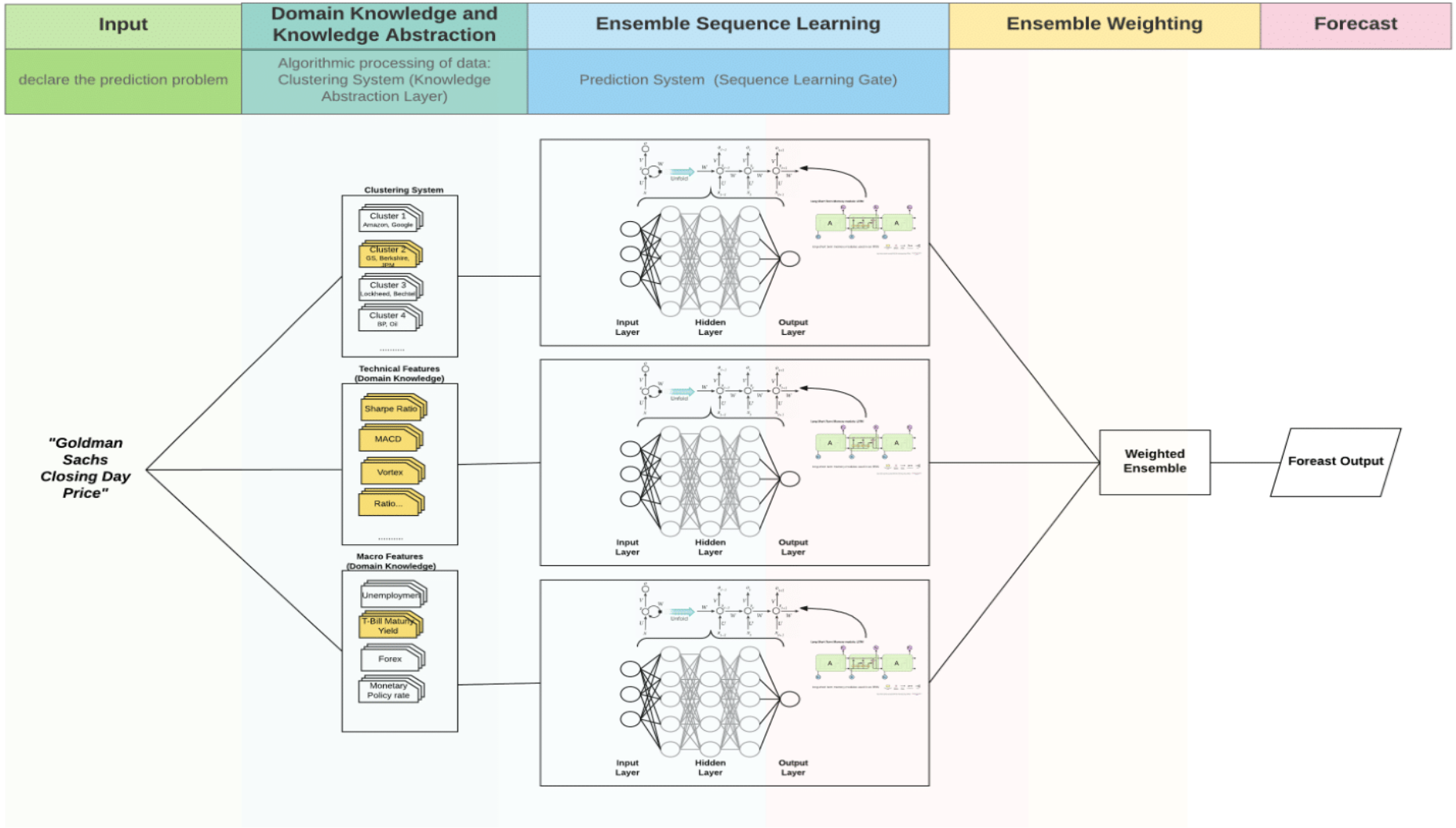


Clustering System



Prediction System

Abstraction of Ensemble Prediction Model Framework



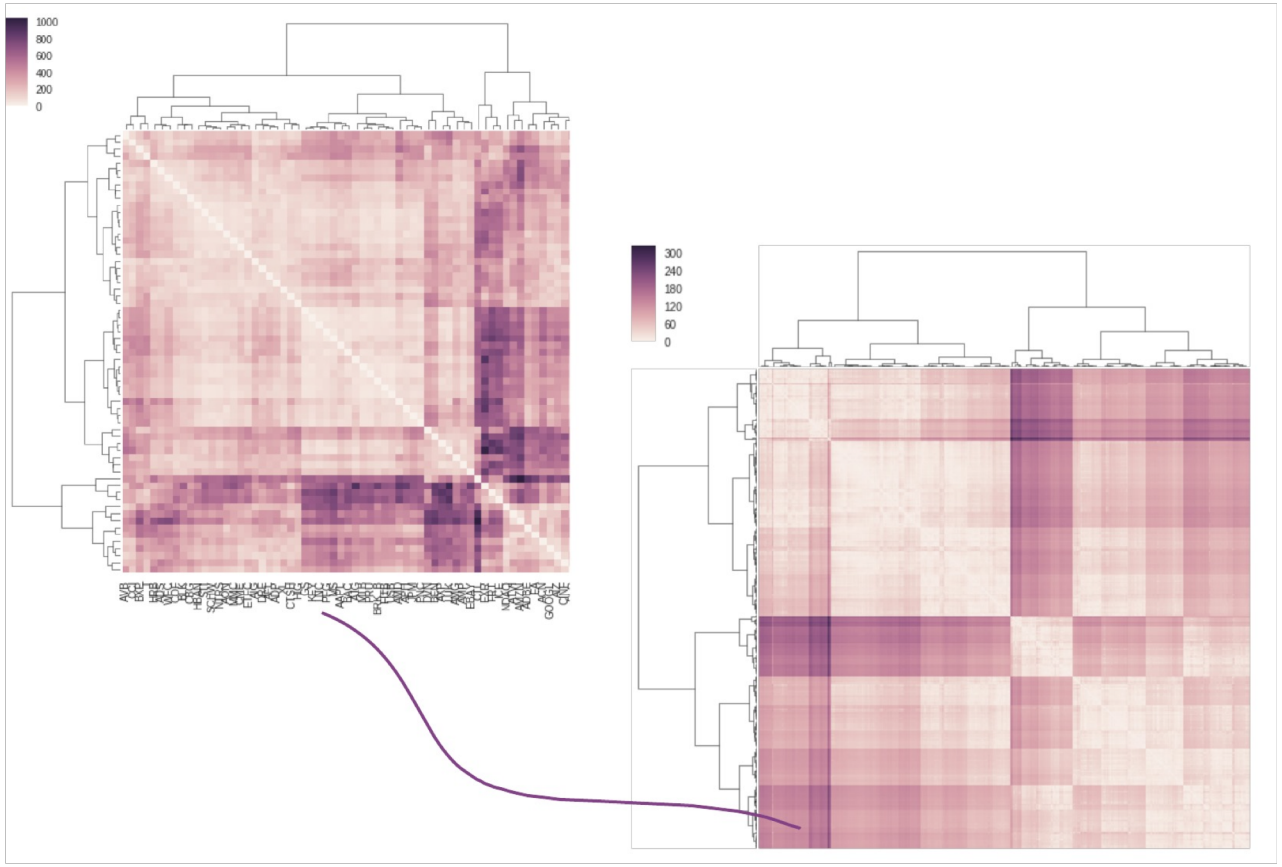
Applying DTW and then Hierarchical Clustering

Calculate DTW distances for Sharpe Ratio of each ticker over 500 days

Make a Matrix of obtained DTW Distance

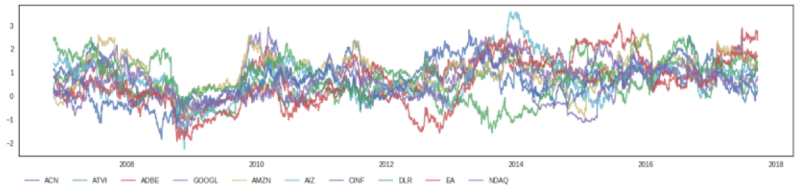
Based on the DTW matrix obtained apply Hierarchical Clustering with linkage='ward'

Here, using distance correlation matrix we create a cluster heat map



Cluster 2, 4, and 5 for selected companies, stock tickers are identified

Cluster 2:

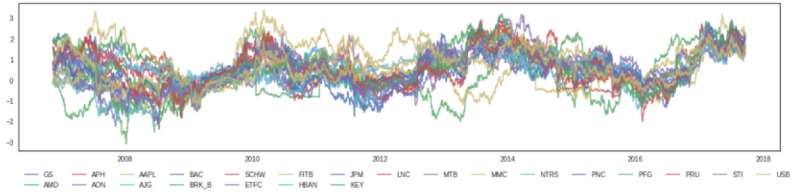


['ACN', 'ATVI', 'ADBE', 'GOOGL', 'AMZN', 'AIZ', 'CINF', 'DLR', 'EA', 'NDAQ']

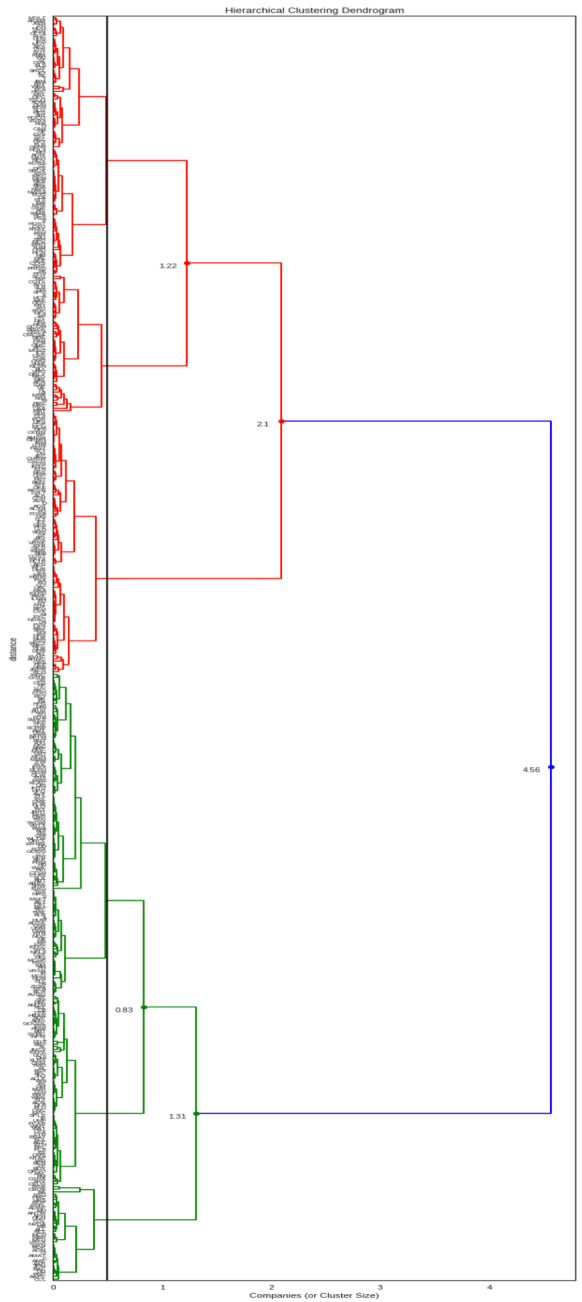
Cluster 4:

['AFL', 'ADS', 'AIG', 'ADP', 'AVB', 'BLK', 'HRB', 'BXP', 'T', 'COF', 'CBG', 'CME', 'CTSH', 'CCI', 'DRE', 'HIG', 'WFC', 'XL']

Cluster 5:

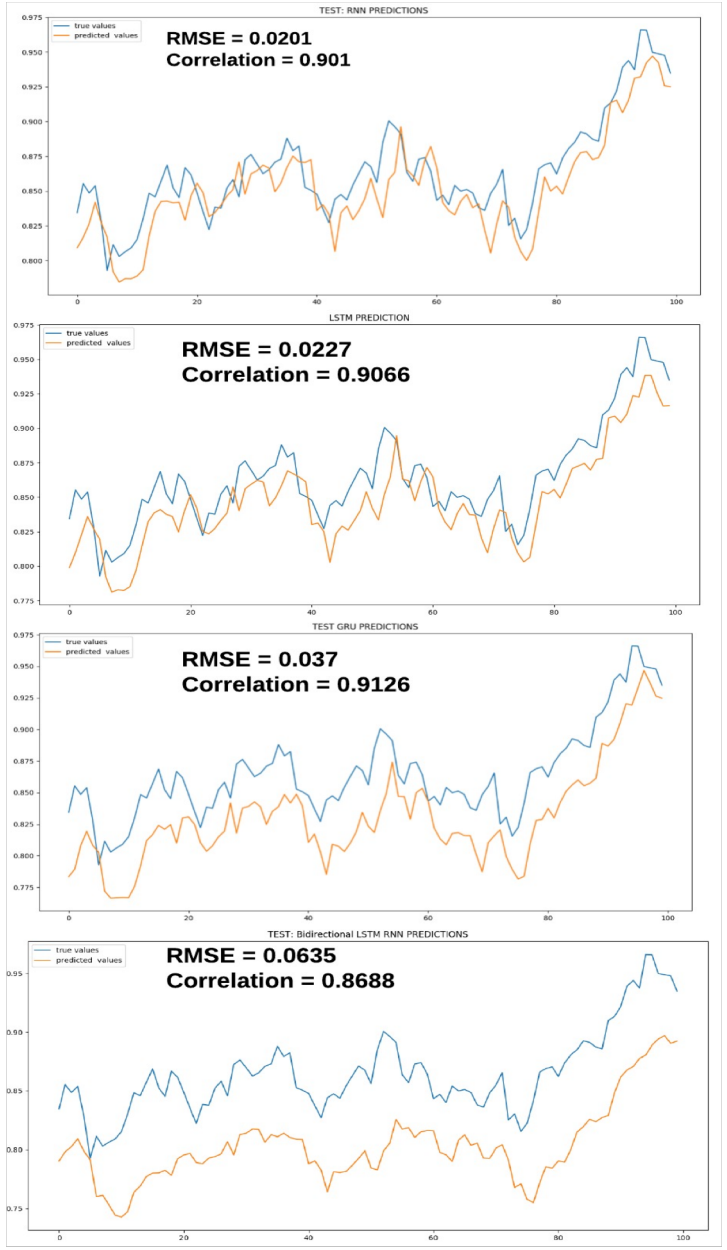


['GS', 'AMD', 'APH', 'AON', 'AAPL', 'AJG', 'BAC', 'BRK_B', 'SCHW', 'ETFC', 'FITB', 'HBAN', 'JPM', 'KEY', 'LNC', 'MTB', 'MMC', 'NTRS', 'PNC', 'PFG', 'PRU', 'STI', 'USB']


Hierarchical Clustering Dendrogram
distance
1.22
2.1
4.56
0.83
1.31
0
1
2
3
4
Companies (or Cluster Size)

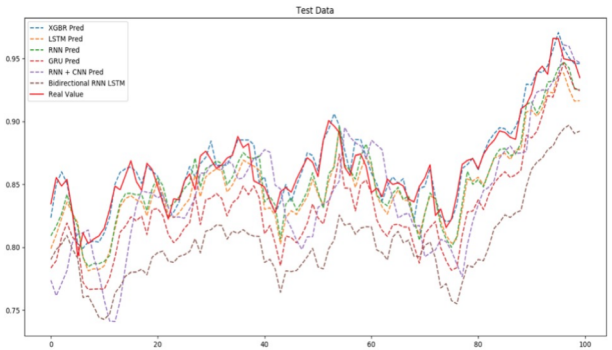
Incorporating different models

Model	Description
LSTM	Long-Short Term Memory
Bi directional LSTM	Bi-directional LSTM
RNN	Recurrent Neural Networks
GRU	Gated Recurrent Unit
GRADIENT BOOSING REGRESSOR	Gradient boosting regressors are a type of inductively generated tree ensemble model. At each step, a new tree is trained against the negative gradient of the loss function, which is analogous to (or identical to, in the case of least-squares error) the residual error.
ENSEMBLE	Ensemble learning is the process by which multiple models, such as classifiers or experts, are strategically generated and combined to solve a particular computational intelligence problem.


TEST: RNN PREDICTIONS
true values
predicted values
RMSE = 0.0201
Correlation = 0.901
LSTM PREDICTION
true values
predicted values
RMSE = 0.0227
Correlation = 0.9066
TEST GRU PREDICTIONS
true values
predicted values
RMSE = 0.037
Correlation = 0.9126
TEST: Bidirectional LSTM RNN PREDICTIONS
true values
predicted values
RMSE = 0.0635
Correlation = 0.8688

Ensembling different models

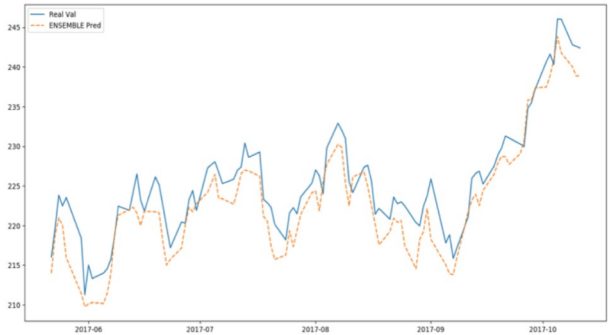
Wealth Generated, S&P500



Comparison between LSTM, GRU, XGBR, RNN for Test Data

Model Parameters and Performance Summary

MODEL	PARAMETERS			RMSE	CORRELATION
	Nodes in each	Hidden Layers	Epochs		
LSTM	256	2	50	0.0152747555996	0.906103767212
Bidirectional LSTM	50	2	50	0.0170483959379	0.906834348288
RNN	256	2	50	0.0201327350773	0.912609408797
RNN + CNN	256	3	50	0.0387413479889	0.738036205928
GRU	256	2	50	0.0374665681995	0.912609408797
Gradient Boosting Regressor (XGBR)	learning_rate: 0.1, max_depth: 6, min_samples_leaf: 5			0.0486210309978	-

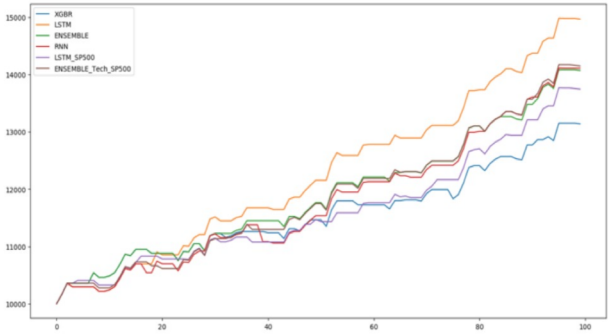


Real Stock prices vs Ensemble (Long Short Term Memory + XG Boost Regression)

Hyper-parameter tuning

INDICATORS	TRAINING DAYS	SEQUENCE	RNN	LSTM	GRU
(Technical Indicators)	250	10	0.069552	0.017427	0.029993
		25	0.054169	0.051353	0.061143
	1000	10	0.033293	0.031885	0.019251
		25	0.032825	0.026986	0.024296

Wealth

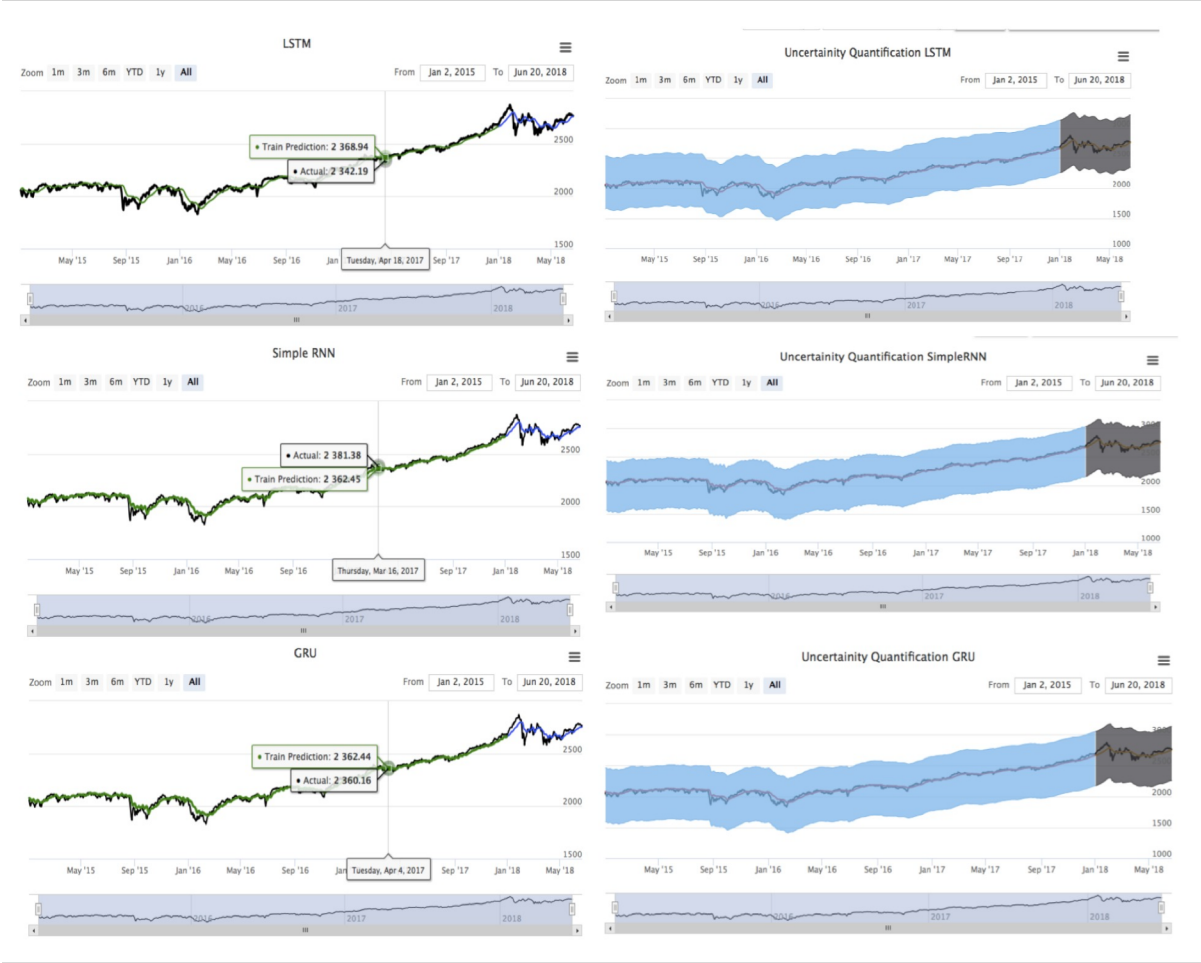


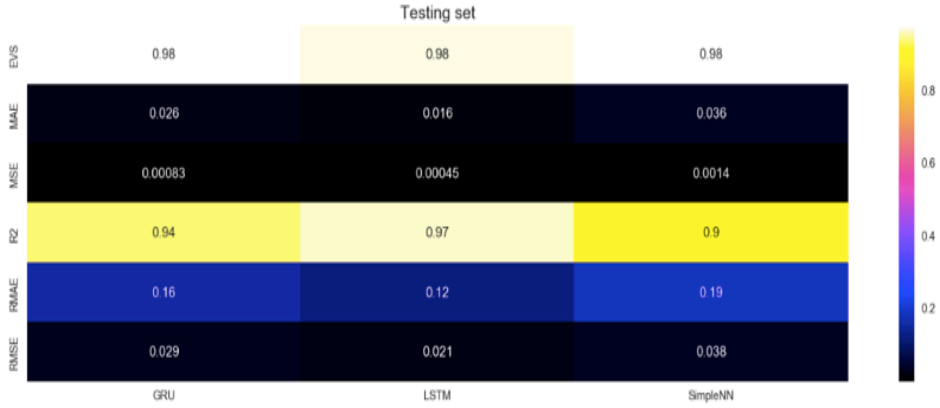
Wealth curve obtained using Closing Price Prediction Multiple Models

Investment Strategy Based Wealth Balance

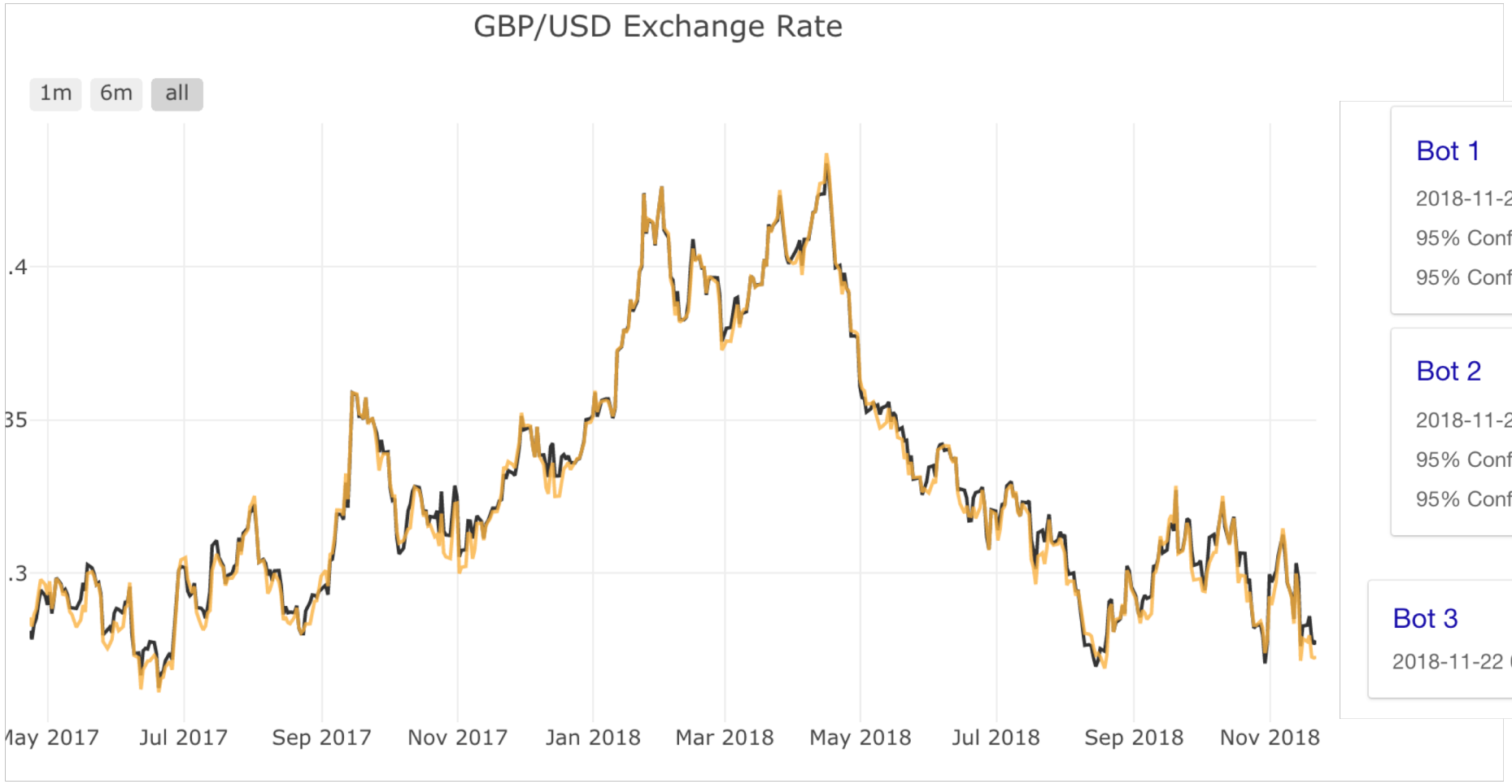
MODEL	Dataset	BALANCE (After 100 Days)
LSTM	SHARPERATIO	10550.840
Gradient Boosting Regressor	Technical Indicators	13140.57
LSTM	Technical Indicators	14968.01
RNN	Technical Indicators	14111.71
LSTM	SP500 Cluster	13745.5
ENSEMBLE 1 (LSTM+XGBR)	SP500 Cluster	14072.96
ENSEMBLE 2 (ENSEMBLE 1 + LSTM)	Technical Indicators, SP 500 Cluster	14151.22

S&P500


LSTM
Train Prediction: 2 368.94
Actual: 2 342.19
Tuesday, Apr 18, 2017
Uncertainity Quantification LSTM
Simple RNN
Actual: 2 381.38
Train Prediction: 2 362.45
Thursday, Mar 16, 2017
Uncertainity Quantification SimpleRNN
GRU
Train Prediction: 2 362.44
Actual: 2 360.16
Tuesday, Apr 4, 2017
Uncertainity Quantification GRU
Zoom 1m 3m 6m YTD 1y All
From Jan 2, 2015 To Jun 20, 2018


Testing set
EVS 0.98 0.98 0.98
MAE 0.026 0.016 0.036
MSE 0.00083 0.00045 0.0014
R2 0.94 0.97 0.9
RMAE 0.16 0.12 0.19
RMSE 0.029 0.021 0.038
GRU
LSTM
SimpleNN
0.8
0.6
0.4
0.2

GBP/USD Exchange Rate Example



Bot 1	1.2931
2018-11-22 02:20 AM UTC	0.0151 (1.1668%) ↑
95% Confidence Higher Bound	1.5301
95% Confidence Lower Bound	1.0561

Bot 2	1.2875
2018-11-22 02:20 AM UTC	0.0095 (0.7394%) ↑
95% Confidence Higher Bound	1.4849
95% Confidence Lower Bound	1.0901

Bot 3	1.2724
2018-11-22 02:20 AM UTC	-0.0056 (0.4425%) ↓

Benchmarking Reliability

Model Reliability Matrix (MRM)

Simulation table to develop trust: which model to trust?

Date	Close	actual_change	Bot1_change	Bot2_change	Bot3_change
2018-09-15 17:03:00	1.3076	0.001	0.0057	0.0042	0.004
2018-09-16 17:03:00	1.3149	0.0073	0.0051	0.003	0.0099
2018-09-17 17:03:00	1.3164	0.0015	-0.0031	-0.0068	0.0041
2018-09-18 17:03:00	1.3138	-0.0026	-0.0054	-0.0103	-0.0002
2018-09-19 17:03:00	1.3271	0.0133	-0.0024	-0.0074	0.0148
2018-09-20 17:03:00	1.3067	-0.0204	-0.0166	-0.0214	-0.0207
2018-09-22	1.3077	0.001	0.0061	0.0018	0.0006

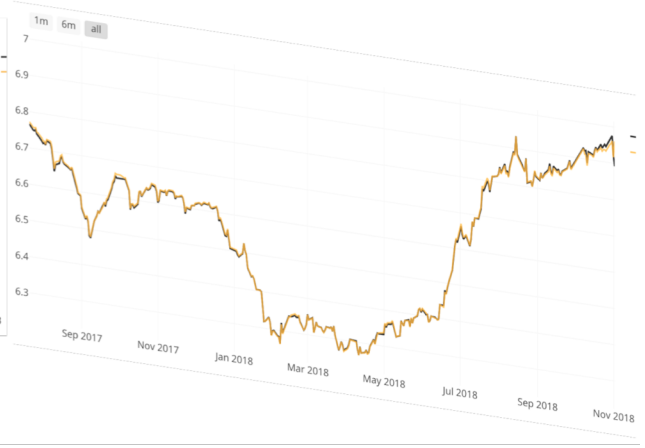
name	Direction %	average change
Bot 1	50	0.0042
Bot 2	58	-0.0032
Bot 3	82	-0.002

Monitoring multiple assets

GBP/USD

USD/EURO

USD/CNY



SP500

London Stock Exchange

Bombay Stock Exchange

Research Contributions

- Promising new application of deep learning to new domains of financial time series predictions
- Structured logic approach for training data in complex environments
- The novel controlled experiment scenario for neural networks, provided new ideas to go beyond just black box – Ensembling, Model Reliability Matrix (MRM), Uncertainty Quantification (UQ), Simulation Tables
- Risk Management Framework: Application of foresight monitoring technologies i.e. integrating of predictions from different domains and time-horizons

Future Research

- Wider testing of models and benchmarking model reliability across domains; automation of hyper-parameter tuning
- Deeper Ensembling techniques with probabilistic and parametric model integration, Causal reasoning; Alert Systems
- Developing standards for data reliability and big data architecture in these domains
- Human Centered AI: Psychology, Cognition, Visualizations, how decision-makers react to forecasts

Conclusions

- Proposed model framework has proven to be scalable across domains
- Uncertainty Quantification and accuracy benchmarking is helpful to bring transparency
- Use cases for informing daily forecast monitor for risk management for trading, policy and mega-project investments
- Applications for geo-politics monitor, macro-and-micro economy – scalability to apply high-and-low frequency data

Thank you!

Selected References

Mishra, S., and B. Ayyub, “Ensemble Deep Learning Framework for Financial Time-Series Predictions using Dynamic Time Warping (DTW) based Hierarchical Clustering Analysis (HCA) and Long-Short Term Memory (LSTM)”, Submitted to Expert Systems and Applications, 2018

Mishra, S., and B. Ayyub, “Predicting Regional Economic Activity using Artificial Intelligence (AI) Methods: Case Study with Indian States”, Forthcoming IASSI QUARTERLY, ISSUE NO.3 (JULY-SEPTEMBER, 2018), INDIAN ASSOCIATION OF SOCIAL SCIENCE INSTITUTIONS

Mishra, S., and B. Ayyub, “Monitoring and Predicting Protests and Mass Movements with Ensemble Deep Learning and the USA as a Case Study”, Submitted to Technology Forecasting and Social Change, 2018
2018

Mishra, S., and B. Ayyub, “Entropy for Quantifying Risk and Uncertainty in Economic Disparity”, Risk Analysis Journal, forthcoming 2018-19

Mishra, S., and B. Ayyub, “Entropy-Based Economic Disparity Profiles: The Tale of Four US Cities Baltimore, Detroit, San Francisco, and Washington, DC”, Manuscript, 2018

Mishra, S., I. Tewari, and S. Toosi, “Economic Complexity and the Globalization of Services”, Submitted to Journal of International Economics and Policy, 2018

Mishra, S., P. Loungani, C. Papageorgiou, and K. Wang, “World Trade in Services: Evidence from a New Dataset”, IMF WP 17/17, Submitted to Journal of Economic Growth, 2018

Mishra, S., I. Tewari, H. Rai, and S. Toosi, “Service Complexity to Inform Diversification Strategies for Managers“, Under Review and Re-Submitting for Journal of Service Research

Mishra, S., “Data Driven Analytical Models for Identification and Predictio of Opportunities and Threats”, PhD Dissertation, University of Maryland College Park

Gonzalez, N., P. Loungani, and S. Mishra, “What Lies Beneath?: A Sub-National Look at Okun’s Law, Open Economies Review Special Issues, forthcoming 2018

Mishra, S., A. Zaccaria, M. Cader, and L. Pietranero, “Integrating services in Economic Fitness Approach”, World Bank Working Policy Paper 2018

Selected References

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. (2017). "How AI Will Change the Way We Make Decisions", Harvard Business Review, Decision Making.

Ayyub, Bilal. (2014). "Risk Analysis in Engineering and Economics", 2014, Taylor and Francis

Ayyub, Bilal, and George Klir. (2006). "Uncertainty Modelling and Analysis in Engineering and the Sciences, Chapman and Hall Press.

Bank of International Settlements (BIS). (2010). "Developments in Modelling Risk Aggregation", Basel Committee on Banking Supervision

Bao W, Yue J, Rao Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLoS ONE 12(7): e0180944. https://doi.org/10.1371/journal.pone.0180944

Deng, X., Pheng, L. S., & Zhao, X. (2014). "Project system vulnerability to political risks in international construction projects: the case of Chinese contractors", *Project Management Journal*, 45(2), 20–33, 2014

Ellis C.J., and J. Fender. (2011). "Information Cascades and Revolutionary Regime Transitions", *The Economic Journal*, Volume 121, Issue 553, Pages 763–792.

Flyvbjerg, B. (2014). "What You Should Know about Megaprojects and Why: An Overview", Project Management Journal, vol. 45, no. 2, April-May, pp. 6-19

Fox, Anthony, Chris Eichelberger, James Hughes, Skylar Lyon. (2013). "Spatio-temporal Indexing in Non-relational Distributed Databases"

Hassan, K., N.C. Maroney, H.M. El-Sady, and A. Telfah. (2003). "Country risk and stock market volatility, predictability, and diversification in the Middle East and Africa", *Economic Systems* 27(1):63-82, DOI: 10.1016/S0939-3625(03)00017-7

Hayashi, Yuko Mizuhara, and Nobuo Suematsu. (2016). "Embedding time series data for classification"

Hubbard, D. (2016). "The Failure of Risk Management: Why It's Broken and How to Fix It", John Wiley & Sons. p. 46).

Selected References

Hohenecker, Patrick, and Thomas Lukasiewicz. (2017). “Deep Learning for Ontology Reasoning”, Artificial Intellignece, arXiv:1705.10342

Holopainen, Markus and Peter Sarlin. (2016). “Toward robust early-warning models: a horse race, ensembles and model uncertainty”, European Central Bank Working Paper No 1900

Neelakantan, Arvind. (2017). “Knowledge Representation and reasoning with deep neural networks”, PhD Dissertation University of Amherst

Liu, Quan, Hui Jiang, Zhen-Hua Ling, Si Wei, and Yu Hu. (2016). “Probabilistic Reasoning via Deep Learning: Neural Association Models”

Locatellia, Giorgio, Giacomo Marianib, Tristano Sainatia, and Marco Grecob. (2017). “Corruption in public projects and megaprojects: There is an elephant in the room!”, International Journal of Project Management, Volume 35, Issue 3, Pages 252-268

Santoro, Adam, David Raposo, David G.T. Barrett, Mateusz Malinowski, Razvan Pascanu, Peter Battaglia, Timothy Lillicrap. (2017). “A simple neural network module for relational reasoning”, Computation and Language, arXiv:1706.01427

Sarlin, P.. (2013a). Decomposing the global financial crisis: A self-organizing time map. Pattern Recognition Letters 34, 1701–1709.

Sarlin, P. (2013. Self-organizing time map: An abstraction of temporal multivariate patterns. Neurocomputing 99(1), 496–508.

Safer, Alan M. and Bogdan M. Wilamowski. (2011). “Using Neural Networks to Predict Abnormal Returns of Quarterly Earnings”

Sakoe,H. and Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. on Acoust., Speech, and Signal Process., ASSP 26, 43-49.

Tsinaslanidis, Prodromos E., Antonis Alexandridis, Achilleas Zapranis , Efstratios Livanis. (2013). “Dynamic Time Warping as a Similarity Measure: Applications in Finance”, Kent Academic Repository

Schmidhuber. J. (2014). Deep Learning in Neural Networks: An Overview. Neural Networks, Volume 61, January 2015, Pages 85-117 (DOI: 10.1016/j.neunet.2014.09.003), published online in 2014.

Shahar, Y. “A framework for knowledge-based temporal abstraction”, Artificial Intelligence Volume 90, Issues 1–2, February 1997, Pages 79-133

Shannon, C., Weaver, W. (1963). A Mathematical Theory of Communication. University of Illinois Press, Champaign.

Srivasta, N., Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov. (1958). “Dropout: A Simple Way to Prevent Neural Networks from Overfitting, Journal of Machine Learning”, Journal of Machine Learning Research 15 (2014) 1929-1958

Tsinaslanidis, Prodromos E., Antonis Alexandridis, Achilleas Zapranis , Efstratios Livanis. (2013). “Dynamic Time Warping as a Similarity Measure: Applications in Finance”, Kent Academic Repository

Wang, Zhiguang, and Weizhong Yan, and Tim Oates, “Time Series Classification from Scratch with Deep Neural Networks: A Strong Baseline”, GE Global Research

World Economic Forum. (2017). Global Risk Report 2017