Stock Price Prediction of Lithium-Ion Battery Manufacturers using Machine Learning and Sentiment Analysis

Predictive Analytics – Abstract Presentation – Summer 2022

NYU

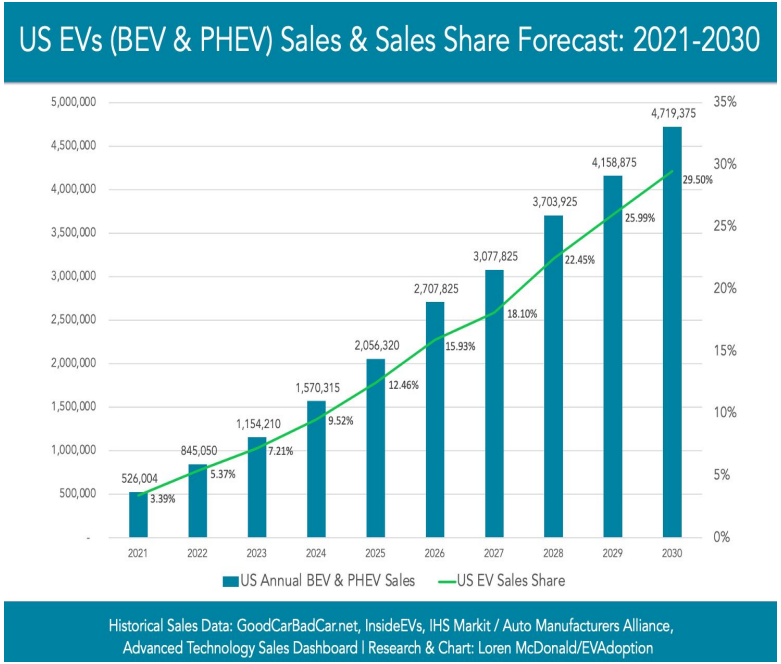
PROJECT MEMBERS:
Anav Prasad
Mitchell Merrick-Thirlway
Sanchit Gaur
Shubham Jha

Introduction

- **Lithium-ion battery** is an energy storage system that's on the verge of a breakthrough in the automobile market triggered by the meteoric growth and expansion of **EVs (Electric Vehicle)** manufacturing companies like **Tesla, Nio and Lucid**.

- Sale projections show that **this growth will exponentially increase**, and **EVs will occupy a larger share** in the automobile market in the **next decade to cater to the rising demand**.

- **Lithium reserves** and the mining industry established around it serves as a **key foundation that fuels** this global shift in **automobile industry** to **green energy sources** from fossil fuels, as concerns about **global warming** grow.

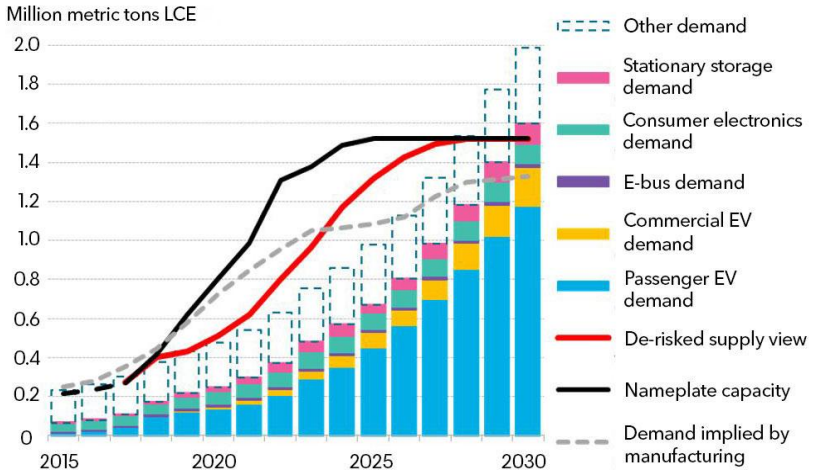


Introduction (Lithium mining)

- Discussions like the supply of extractable lithium deposits, skyrocketing prices of ore, the perception of **worldwide lithium shortage, geopolitical scenarios** governing the trade and the **environmental impact** of mining itself have **reverberations** across the whole supply chain of these **interlinked industries**.

- The **high demand for Li-Ion batteries (and consequently, Lithium ore)** will grow with the **increase in demand** for EVs. The strain on Lithium mining companies (and their **profits)** are **tied to the shifting trend** in automobile market.

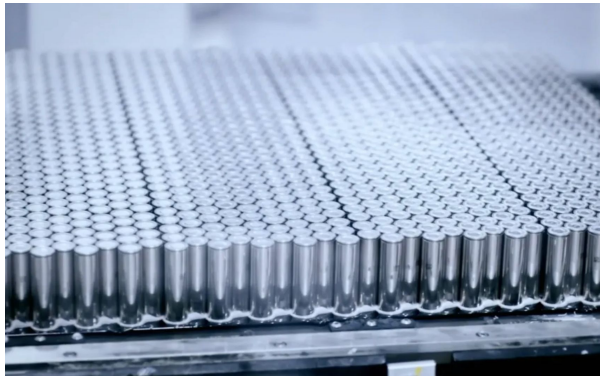
Figure 1: Global lithium supply and demand forecast, comparing methodologies



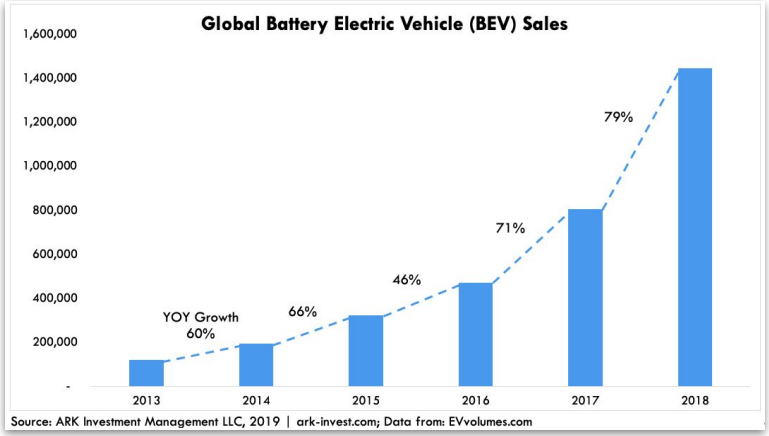
Source: BloombergNEF, Avicenne.

Hypothesis

Since **Lithium** is one of the **major raw materials** for the production of **Li-Ion batteries**, **we hypothesize** that the performance of **mining companies** will have a **direct impact** on the performance of **Li-Ion battery manufacturers' stock price**.



Abstract



With the adoption of **electric vehicles** showing no signs of slowing down, the **dependency on lithium-ion batteries** is ever-increasing. **This project attempts to predict the stock price of lithium-ion battery manufacturers by the stock price of lithium mining companies.** The data sets used includes historical **financial data** of the **manufacturing and mining companies**. We also employ **sentiment analysis** of **financial news** outlets' headlines and press releases related to lithium mining from reputable media outlets. As this is a time series predictive problem and Box–Jenkins [1] suggests **Autoregressive Integrated Moving Average (ARIMA)** is fruitful when working in this area, ARIMA is used, as well as other machine learning algorithms to increase the accuracy of the model.

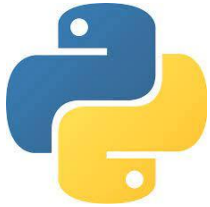
[1] G.P. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, Holden-day Inc., San Francisco, CA, 1976.

Tools


spaCy


scikit
learn


pandas




Transformers


RAPIDMINER


plotly

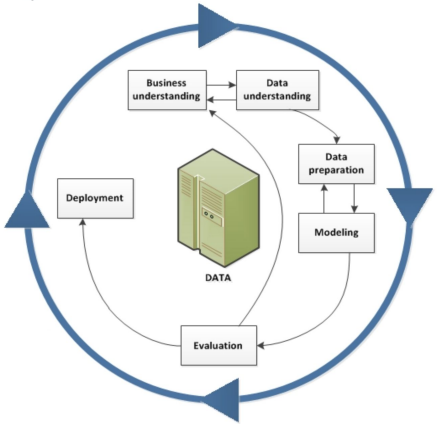
Data Sources

Lithium Mining Companies

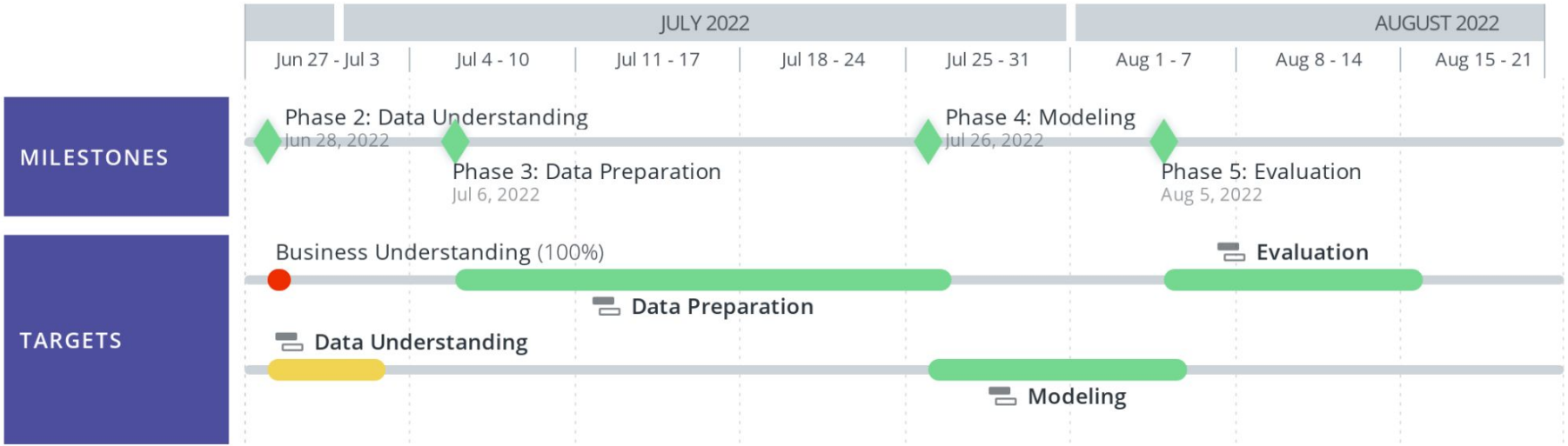
Li-Ion Battery Manufacturers

Roadmap

	SCHEDULED	IN PROGRESS	COMPLETED
TARGETS	**Data Preparation** Data Cleaning Data Formatting Data Integration	**Data Understanding** Data Collection Data Exploration Data Verification	**Business Understanding** 100%
	Modeling Model Selection Model Testing Model Building Model Assessment		
	Evaluation Result Evaluation Process Review		



Timeline


JULY 2022
AUGUST 2022
Jun 27 - Jul 3
Jul 4 - 10
Jul 11 - 17
Jul 18 - 24
Jul 25 - 31
Aug 1 - 7
Aug 8 - 14
Aug 15 - 21
MILESTONES
Phase 2: Data Understanding
Jun 28, 2022
Phase 3: Data Preparation
Jul 6, 2022
Phase 4: Modeling
Jul 26, 2022
Phase 5: Evaluation
Aug 5, 2022
TARGETS
Business Understanding (100%)
Data Preparation
Evaluation
Data Understanding
Modeling

Papers (1)

Study of Effectiveness of Time Series Modeling (ARIMA) in Forecasting Stock Prices

– Prapanna Mondal, Labani Shit and Saptarsi Goswami

This paper was selected as **ARIMA** (Autoregressive Integrated Moving Average) is widely **used to predict linear time series data**. In this paper, the writers investigated the **effectiveness of the ARIMA model in predicting stock prices** of seven sectors of the Indian National Stock Exchange, using twenty-three months' worth of historical data. They split this data into **three different time periods** (six, twelve, and eighteen months) to test the **effect of sample size** in the ARIMA model on price prediction. Using **Akaike Information Criteria (AIC)** to select the best parameters for the ARIMA model, they found, for all seven sectors, **the accuracy of ARIMA in predicting stock prices was above 85%**.

As a result of reading this paper, **we will further investigate AIC** in selecting the best model to use in our models; an extension of ARIMA using **fuzzy regression (FARIMA)**; and **ARIMAX**, which includes other time series as input variables.

Papers (2)

Stock Price Prediction Using News Sentiment Analysis

– Saloni Mohan, Sahitya Mullapudi, Sudheer Sammeta, Parag Vijayvergia, David C. Anastasiu

The research paper Stock Price Prediction using News Sentiment Analysis by Saloni Mohan et al. discusses the strong correlation between a **company's stock price movement** and **news articles published** about it and assesses the viability of **real-time stock price prediction**.

Their work improves upon the traditional methodologies used in contemporary studies (that have utilized SVMs, naive **Bayes regression**, etc.) by incorporating a large volume of time series data (**5-year stock prices of S&P500** companies) in addition to collecting a larger dataset of **265,000 news articles** to analyze through cloud computing, resulting in better accuracy.

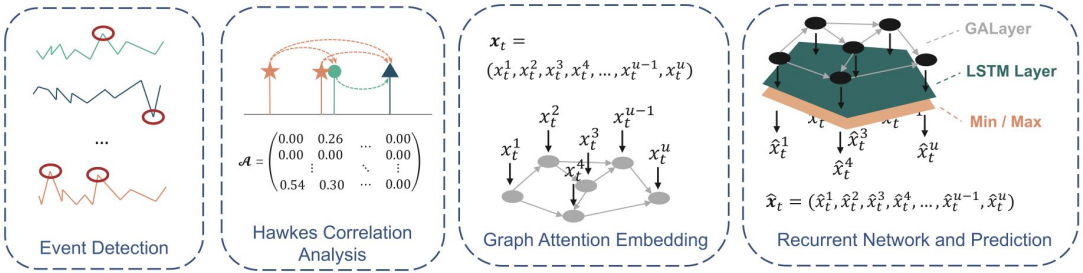
They have utilized **deep learning** time series prediction models based on **ARIMA**, **Facebook Prophet**, and **RNN Long Short Term Memory** and found that the **Recurrent Neural Network**-based approach performed better (especially for stable stocks) and found a correlation between the **text information** and **stock price movement**.

Papers (3)

Graph-based stock correlation and prediction for high-frequency trading systems

– Tao Yina, Chenzhengyi Liub, Fangyu Dingb, Ziming Fengc, Bo Yuand, Ning Zhangb

- Propose GALSTM (Graph-Attention Long Short-Term Memory)
  - A graph based model capable of learning the correlations between different stocks
  - Use a multi-dimensional Hawkes process to initialize the correlation graph
  - The Graph Attention layer learns the correlations between different stocks, and the LSTM learns the temporal nature of the stock ticker.
- Introduce the concept of FairPrice, and perform experiments to empirically determine the best formulation.
- Perform experiments on all Chinese A-share stock data over a period of 3 months
- Achieve an annual return rate of 44.71%, daily std-dev 0.42% (over 3 months)

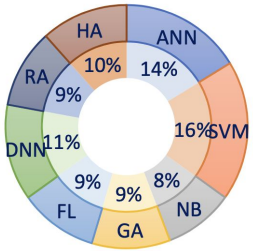


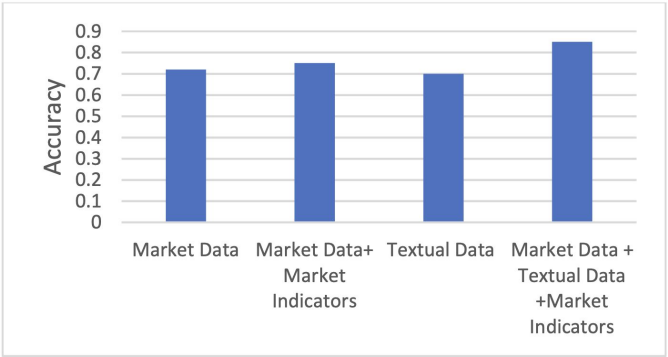
Papers (4)

Stock Market Prediction Using Machine Learning: A Decade Survey on Methodologies, Recent Developments, and Future Directions

– Nusrat Rouf, Majid Bashir Malik, Tasleem Arif, Sparsh Sharma, Saurabh Singh, Satyabrata Aich and Hee-Cheol Kim

- Traditional Stock Market Prediction (SMP) methods: Fundamental and Technical Analysis.
- Modern Approaches using Machine Learning: SVM, kNN, ANN, Decision Trees, Fuzzy Time Series, Evolutionary Algorithms.
- Benefits of alternative data sources: Sentiment Analysis
- Evaluation Metrics: MSE, AUC, AIC, R2, MAE, MAPE
- Most Popular Approach: SVM
- Most Accurate Approaches: ANN, DNN





Papers (5)

Comparison of Stock Price Prediction Models Using Pre-trained Neural Networks

– C. Anand

- …

References

- https://www.statista.com/chart/23808/lithium-ion-battery-demand/
- https://www.statista.com/statistics/1247301/build-out-of-lithium-ion-battery-factories/
- https://www.mckinsey.com/industries/metals-and-mining/our-insights/lithium-mining-how-new-production-technologies-could-fuel-the-global-ev-revolution
- https://www.greenbiz.com/article/lithium-mining-booming-heres-how-manage-its-impact
- https://www.researchgate.net/publication/276197260_Study_of_Effectiveness_of_Time_Series_Modeling_Arima_in_Forecasting_Stock_Prices
- *https://www.researchgate.net/publication/336087787_Stock_Price_Prediction_Using_News_Sentiment_Analysis*
- https://www.researchgate.net/publication/356008402_Stock_Market_Prediction_Using_Machine_Learning_Techniques_A_Decade_Survey_on_Methodologies_Recent_Developments_and_Future_Directions