

SMARTLIST

Used Car Price Range Prediction

By Anthony Kwok
30 Nov 2023

[\[Source Code\]](#)



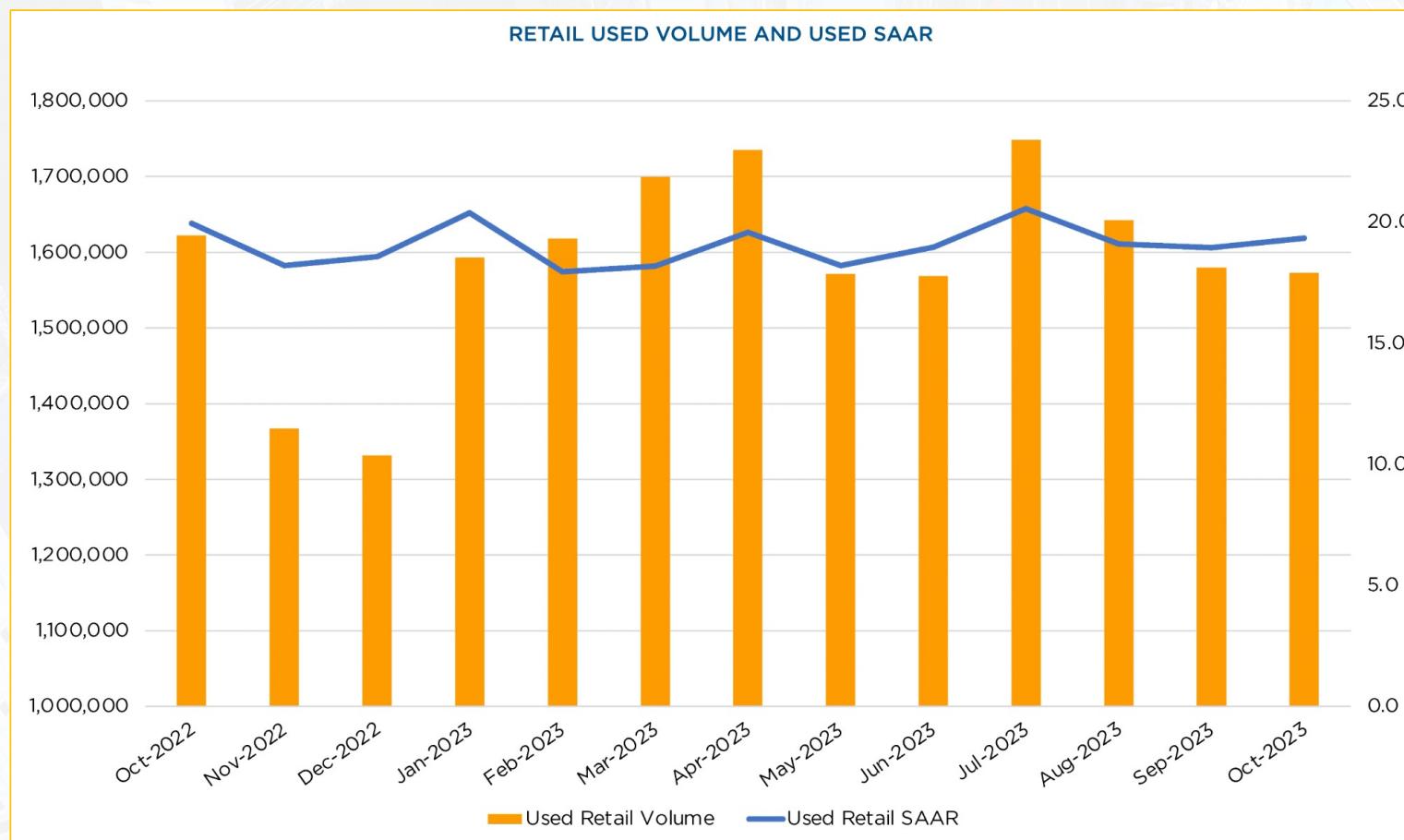
Problem Statement

In average, It takes 4 weeks to sell a used car.

How can we speed up the buying/selling process? To help both buyer and seller to achieve their goal?



Market is Growing!



Monthly Volume

The monthly volume of retail used vehicle in Oct 2023 is around 1.55M ~ 1.75M units.

*Source: <https://www.coxautoinc.com/market-insights/estimated-monthly-used-vehicle-saar-and-volume/>

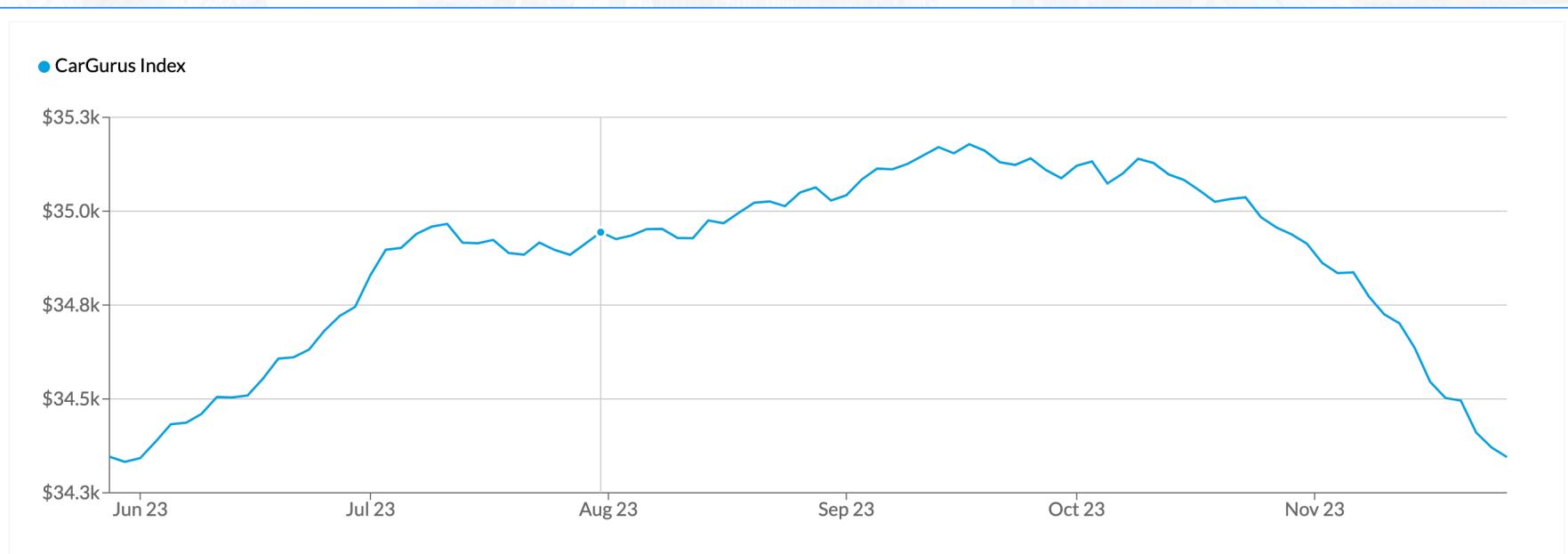
Market is Growing!



Average Used Vehicle Price

According to CarGurus Index, The average used vehicle price in Oct 2023 is around USD 35,000.

**Total Sales Volume
57.8 Billions**



*Source: <https://www.cargurus.ca/research/price-trends>



Who Cares?



Buyer

- Price Uncertainty
- Worry about Overpayment

Dealer

- Hard to manage expectation on both sides
- Long trading time

Seller

- Lack of Knowledge about Vehicle Pricing
- Missed Opportunity



What we need is

Buyer

- Informed Decision Making
- Financial Budgeting
- Accurate Evaluation

Seller

- Reference Price for Listing
- Comparative Analysis

Accurate
Reference
Pricing



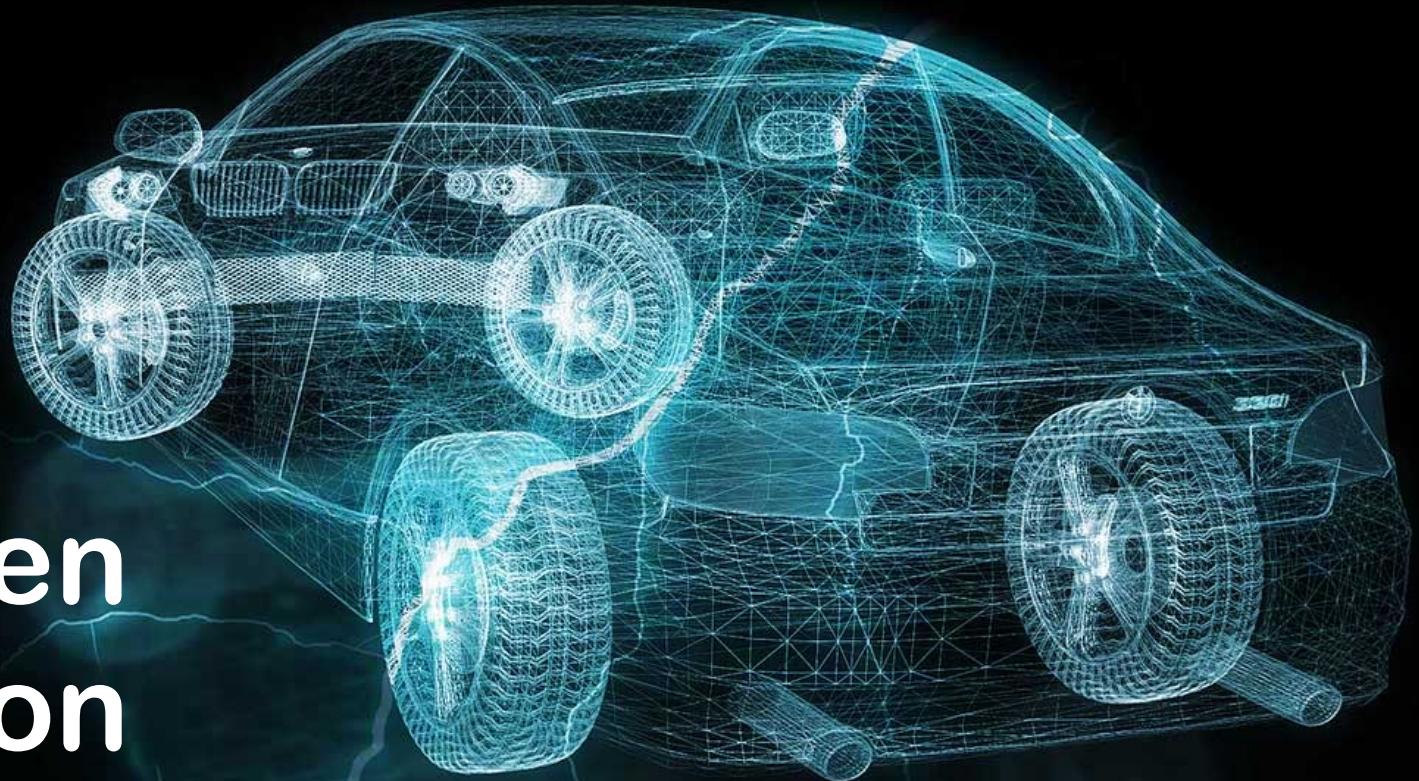
Market Efficiency

- Smoothen the Car Selling Process

Dealer

- Speed Up Matching
- Managing Expectations

Data-Driven Solution



Data-Driven Solution



Specification

Machine Learning

Price Range

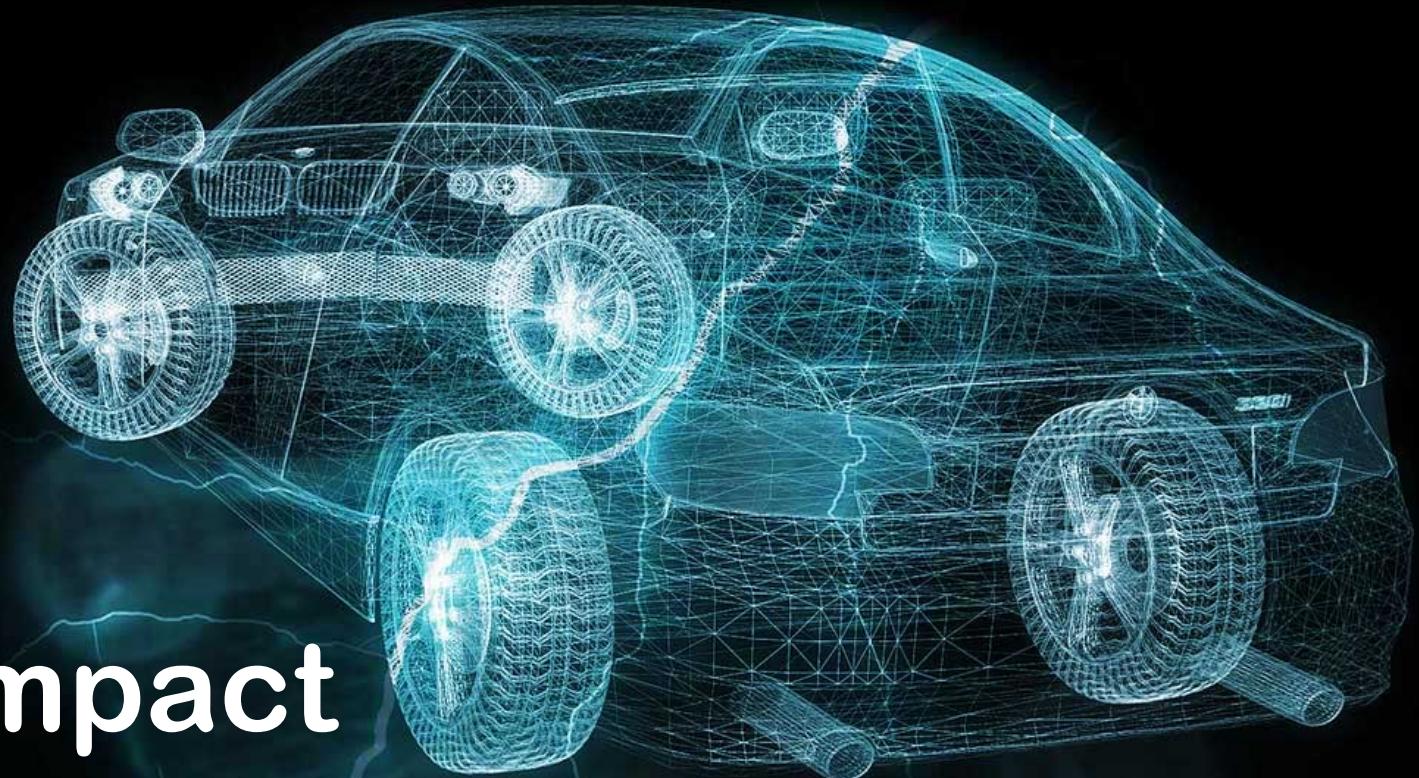
History

AI Solution

Speed Up



Potential Impact





Business Impact



**Speed up matching
time by 25%**



Market Competitors



1. eBay Motors

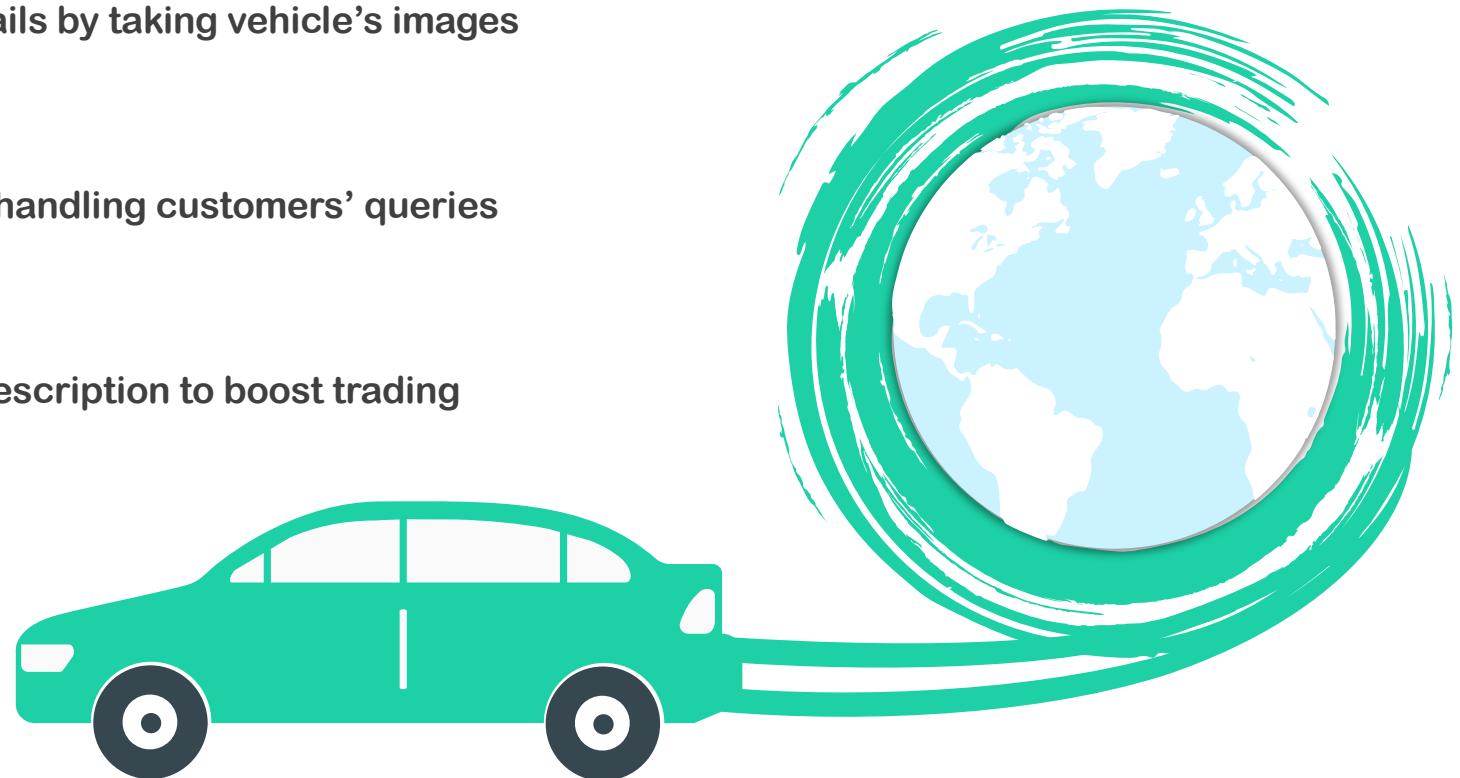
Auto-fill vehicle details by taking vehicle's images

2. Fullpath

Adopt ChatGPT for handling customers' queries

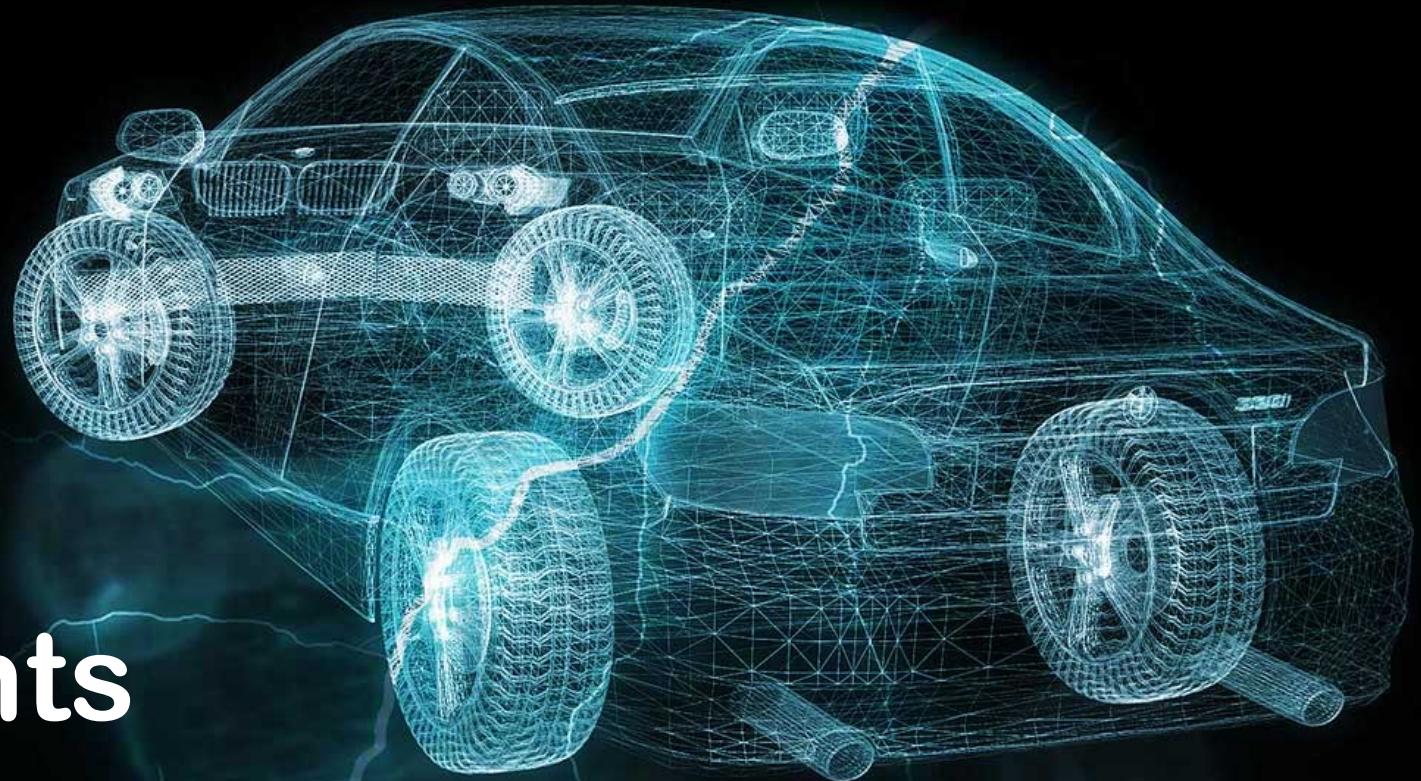
3. AutoRevo

Use AI-generated description to boost trading





Key Insights



Business Insights

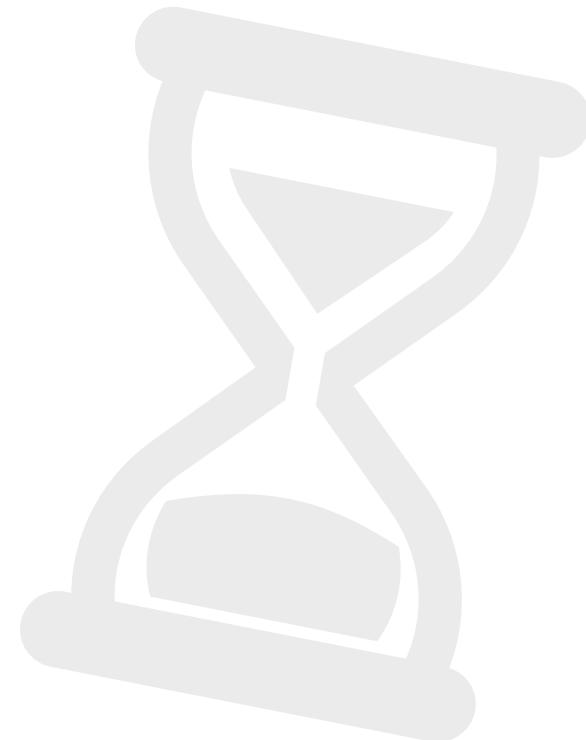


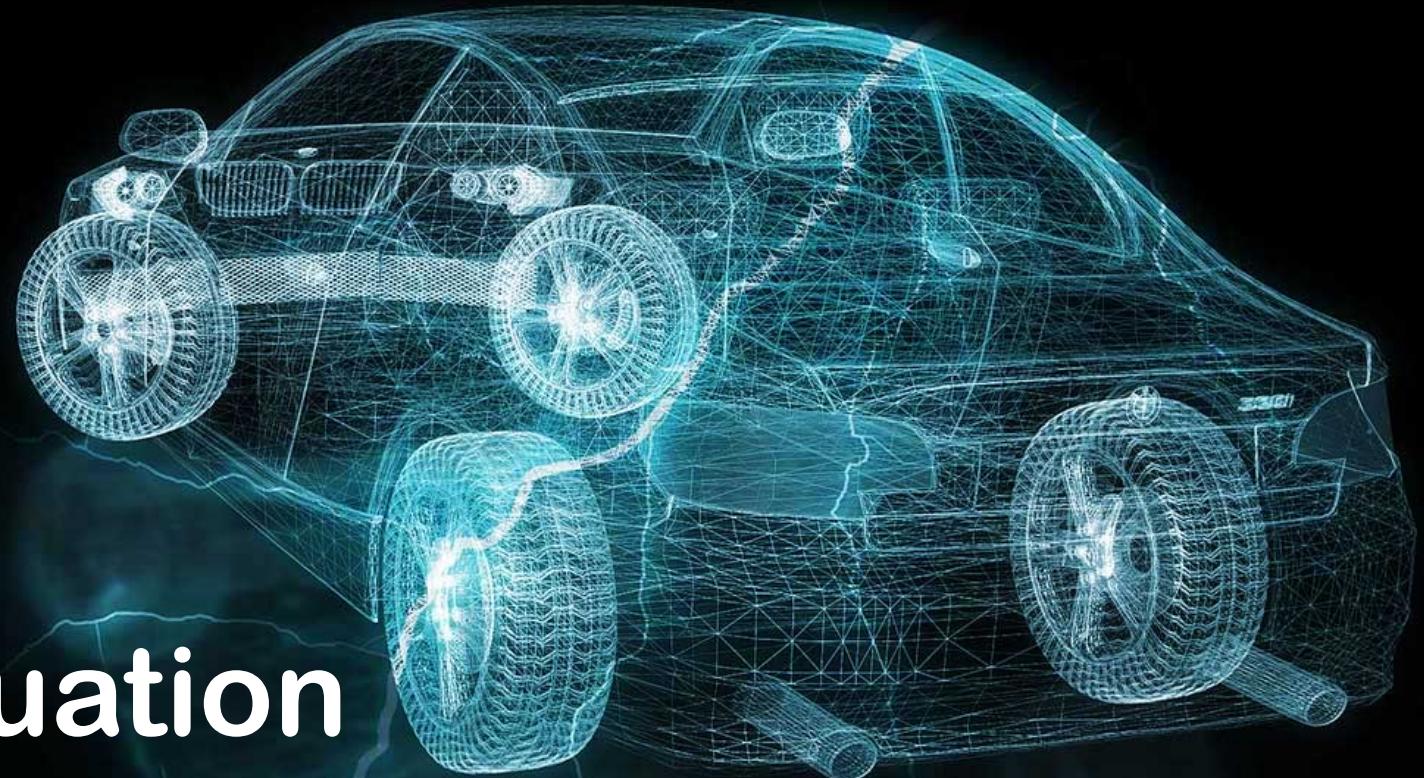
1 | Factor Analysis

Discover patterns in data

2 | Factor-Price Analysis

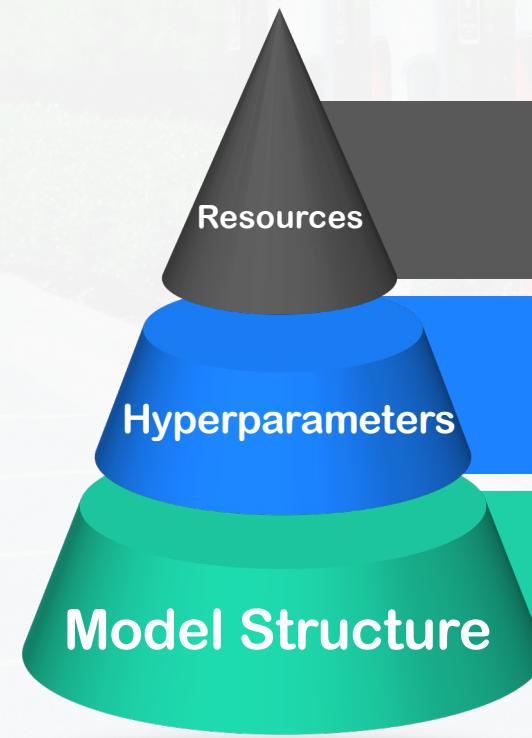
Explore relationships between factors and pricing





Model Evaluation

Model Evaluation Criteria



Balance between performance and resources.

Try different settings of the model.

Different model structures give different results.

Model Evaluation



Model Performance (Full Set)

Model	Training Time	Training F1 Score	Testing F1 Score
Adaptive Boosting (AdaBoost)	3h 14m 20s	86.95%	87.06%
eXtreme Gradient Boosting (XGBoost)	34m 43s	87.32%	87.46%
Random Forest (RF)	20m 47s	87.30%	87.47%

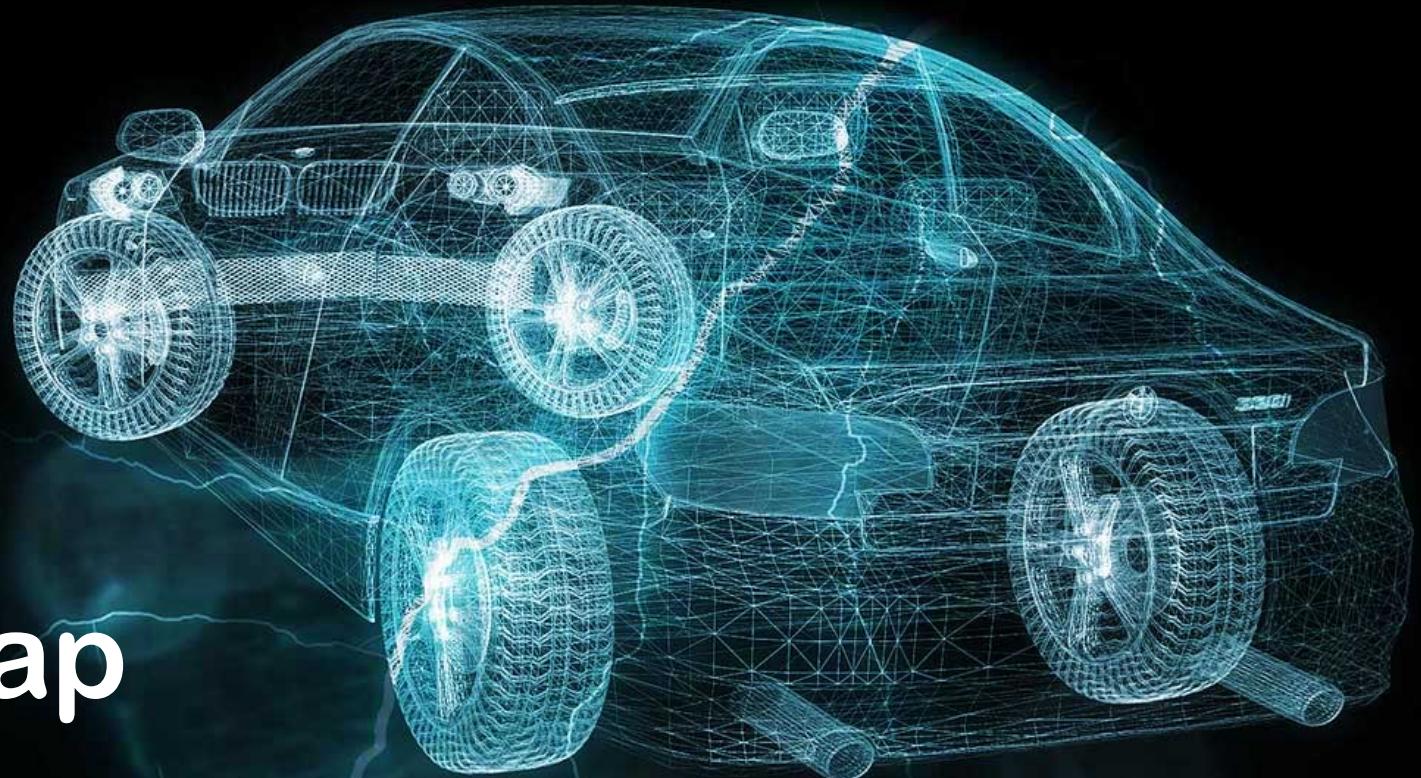


Demo Time





Roadmap

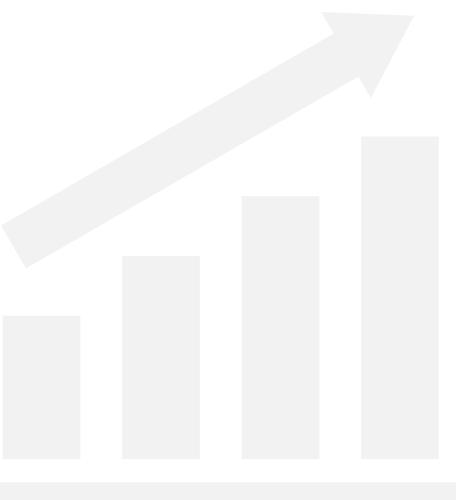


Roadmap



Sprint 0

- Problem Space Definition
- Scope Definition
- Dataset Research



Sprint 1

- Data Cleaning
- Exploratory Data Analysis
- Statistical Analysis



Sprint 2

- Data Preprocessing
- Baseline Model Training



Future

- Introduction of external factor
- Re-define price range (e.g. 5K per range)
- Continually re-train model
- Challenge Ourselves → Beat our RF model

Sprint 3

- Advance Modelling
- Web UI Integration
- Product Delivery



THANK YOU

Follow Me Now!

