Profit Predictor of Companies using LSTM

PATTERN RECOGNITION AND MACHINE LEARNING - Master's in Automatic Control and Robotics

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

UPC

Joan Font Casadevall

Ivan Quirante González

Laura Cervera Sancho

Table of contents

- Data Extraction and management pipeline
- Data Analysis and Exploration
- Baseline Models: SVR and RFR
- LSTM Model and Training
- Evaluation and Results
- Global Comparison
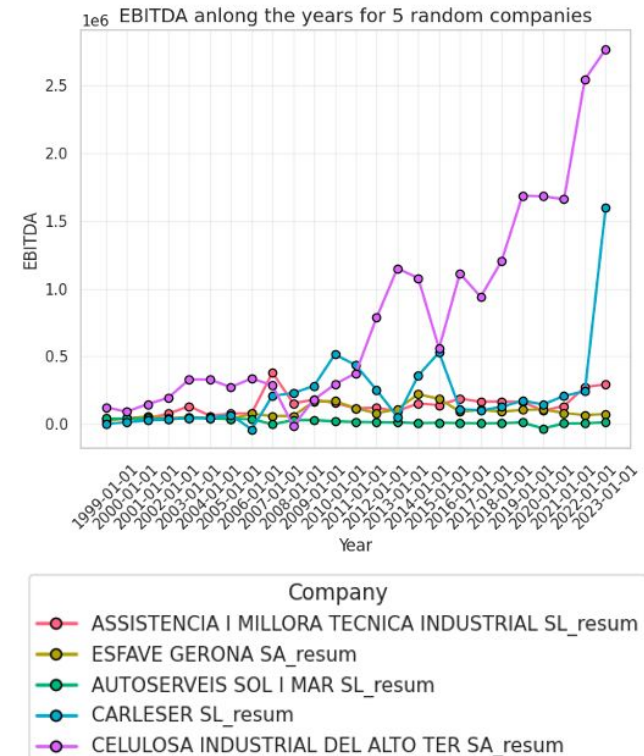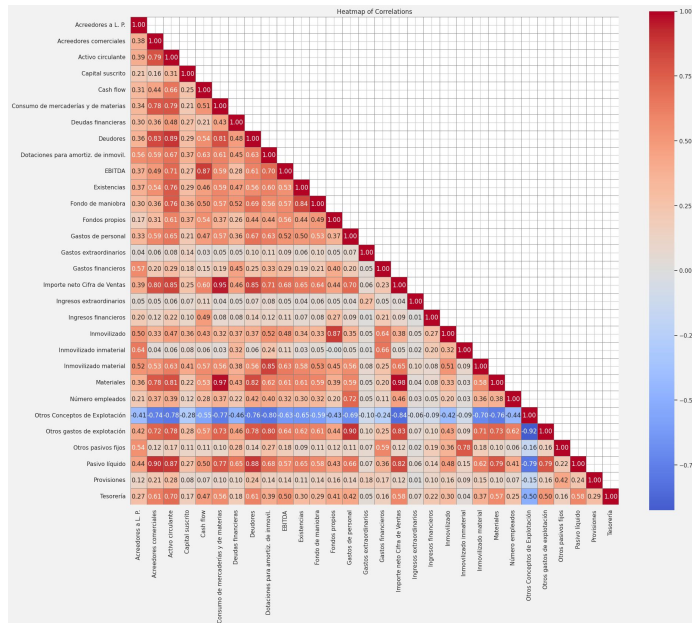- Conclusions

Data Extraction and Management pipeline

-> Source: SABI platform.

-> Data scraping using humanlike interaction techniques.

-> Scripts to convert .xls to csv clean and imputation were implemented.
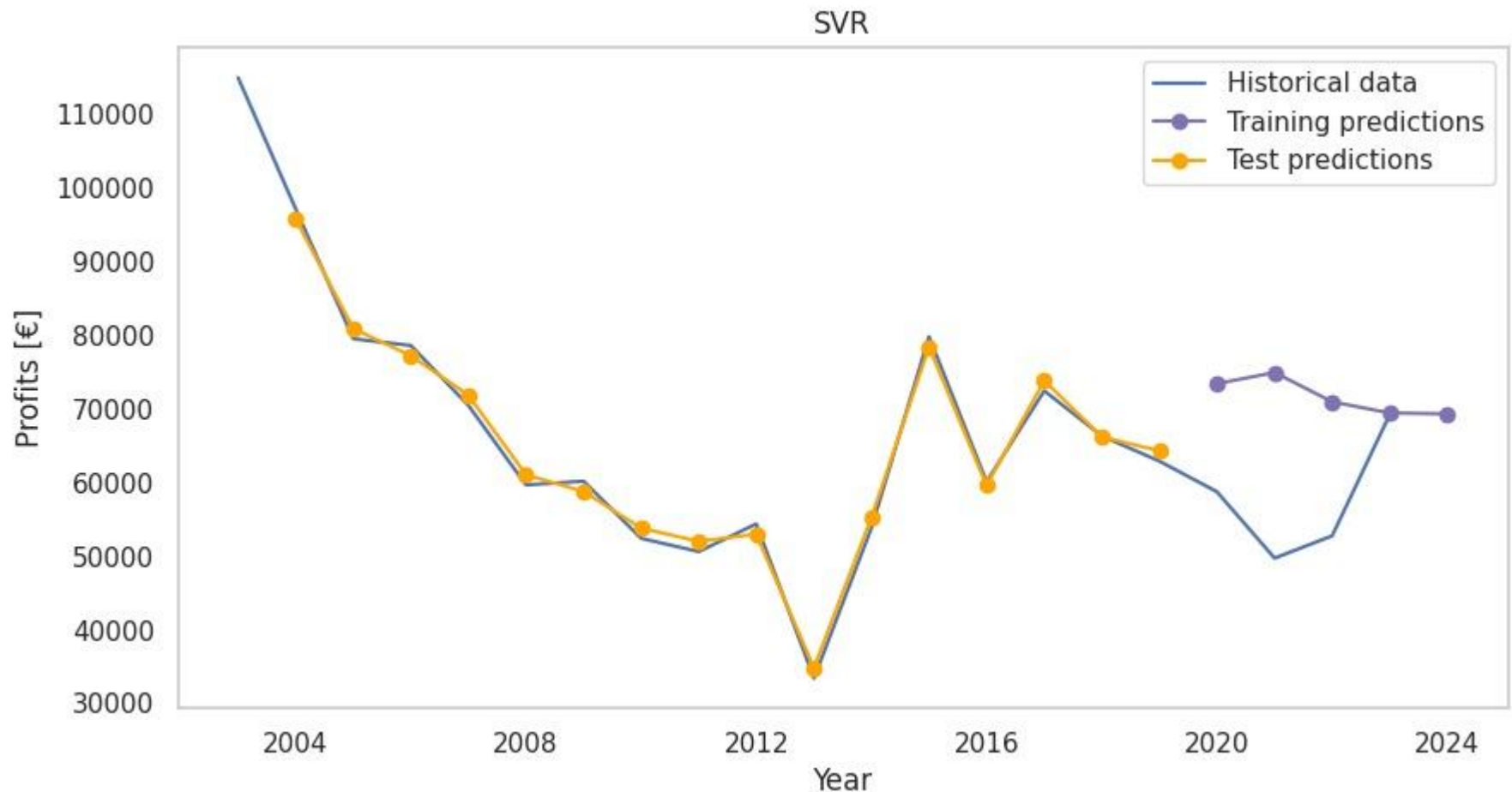
After scraping, filtering and cleaning, the resulting dataset comprises 2,159 unique companies with complete yearly financial data spanning from 2000 to 2023.
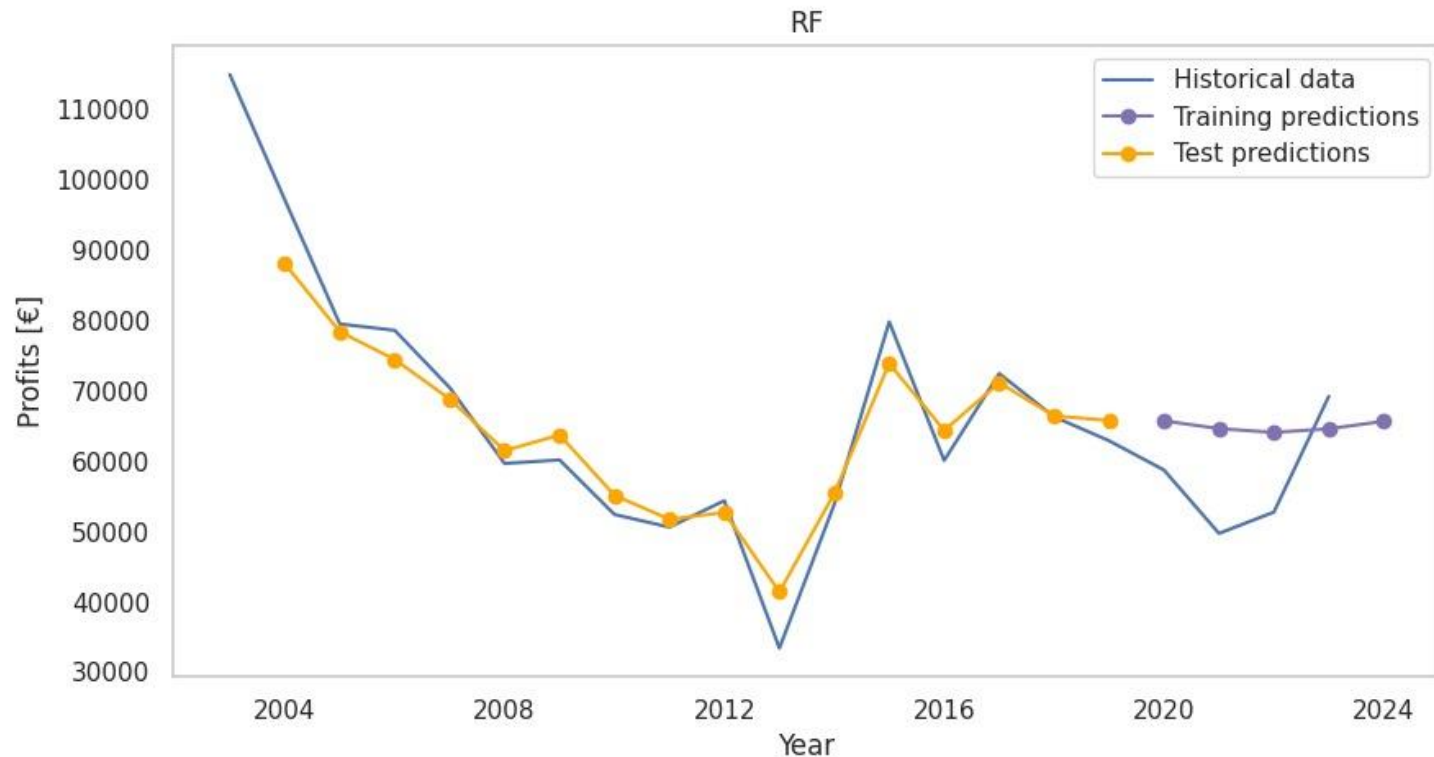
Data Analysis and Exploration

Heatmap of Correlations

EBITDA anlong the years for 5 random companies

SVR Predictions

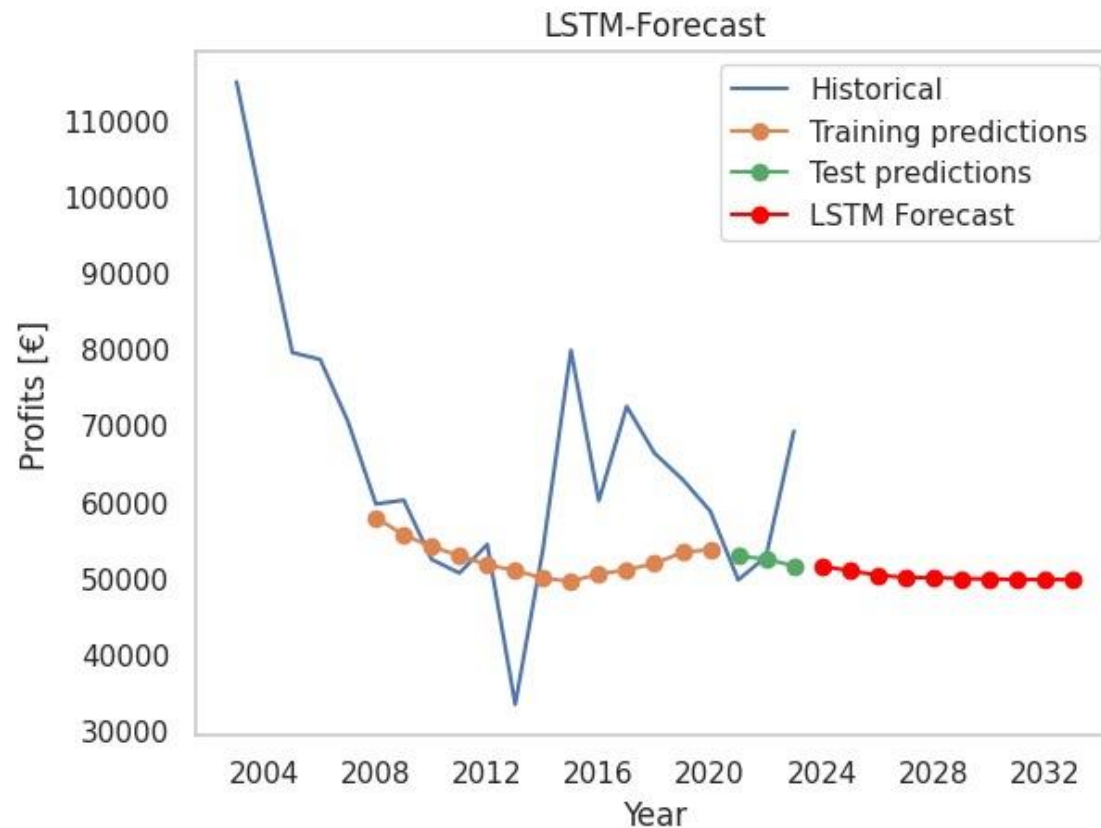- SVR model predictions for a single company.

Random Forest Predictions

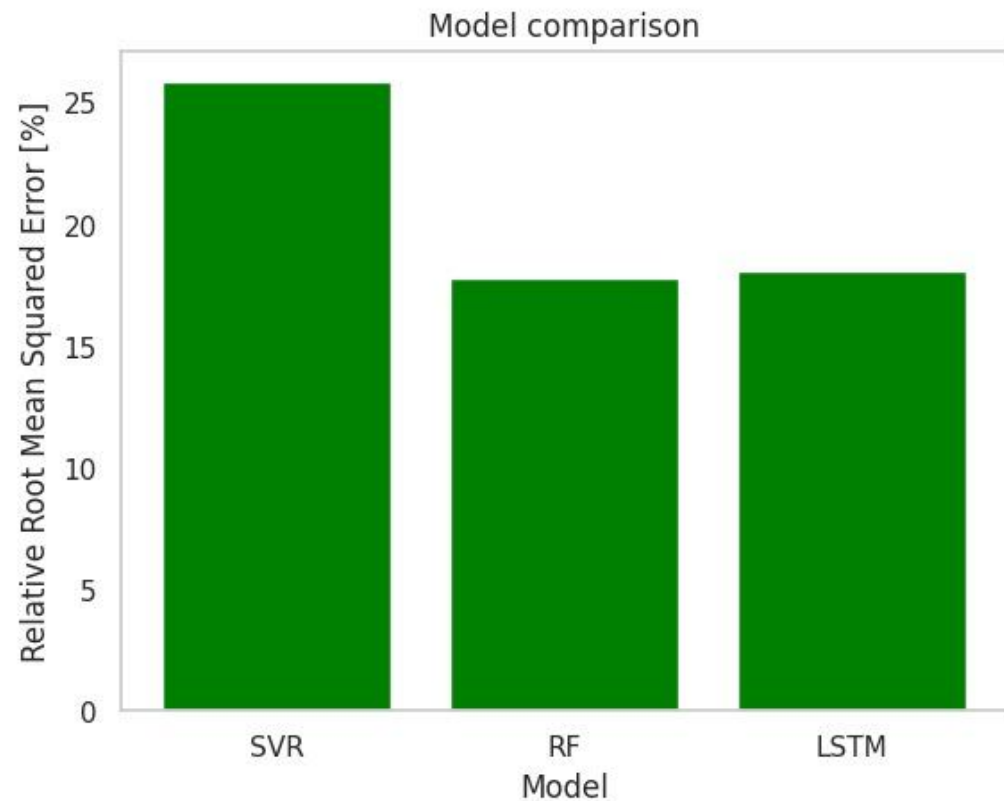- Random Forest predictions for a single company.

LSTM Prediction

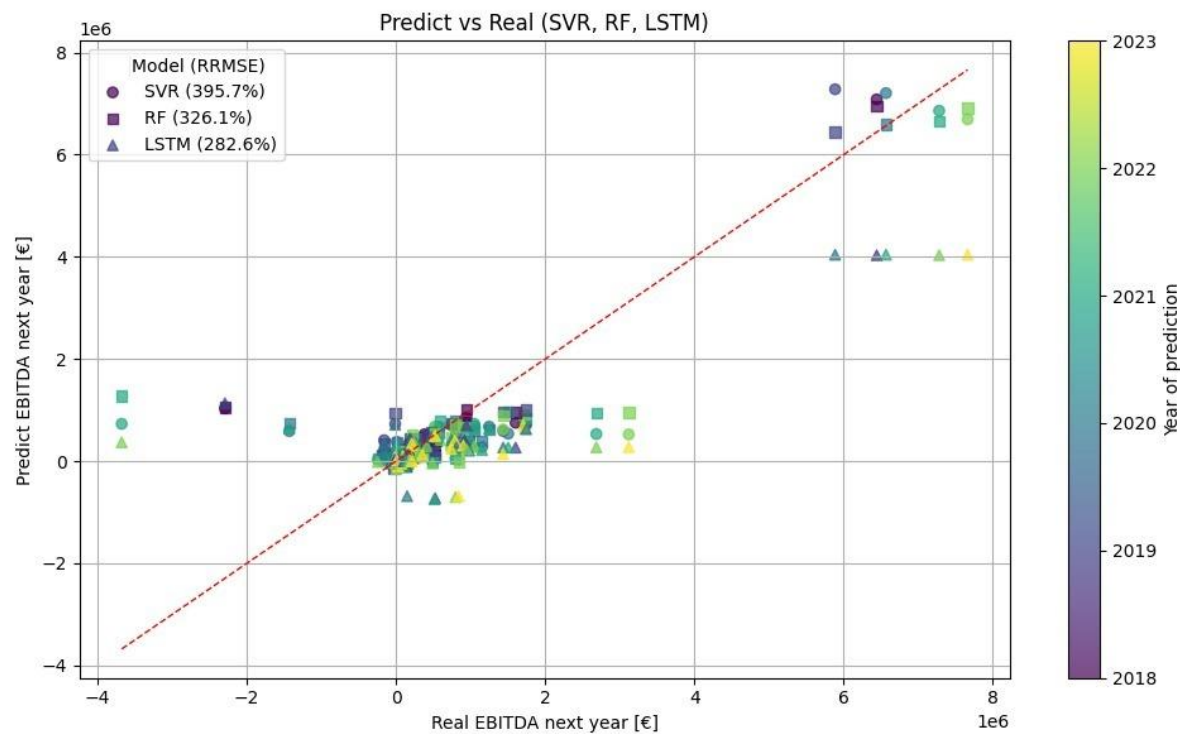- LSTM model predictions for a single company.

RRMSE Comparison

- Relative Root Mean Squared Error (RRMSE) for SVR, RF, and LSTM.

Model comparison

Global Comparison

- Predicted vs Real EBITDA values across all companies.

Conclusions

- High variability and lack of homogeneity across firms

- Three modeling approaches were evaluated

- None of the models achieved accuracy levels suitable for real-world deployment