

INNOVATION

1)Prophet

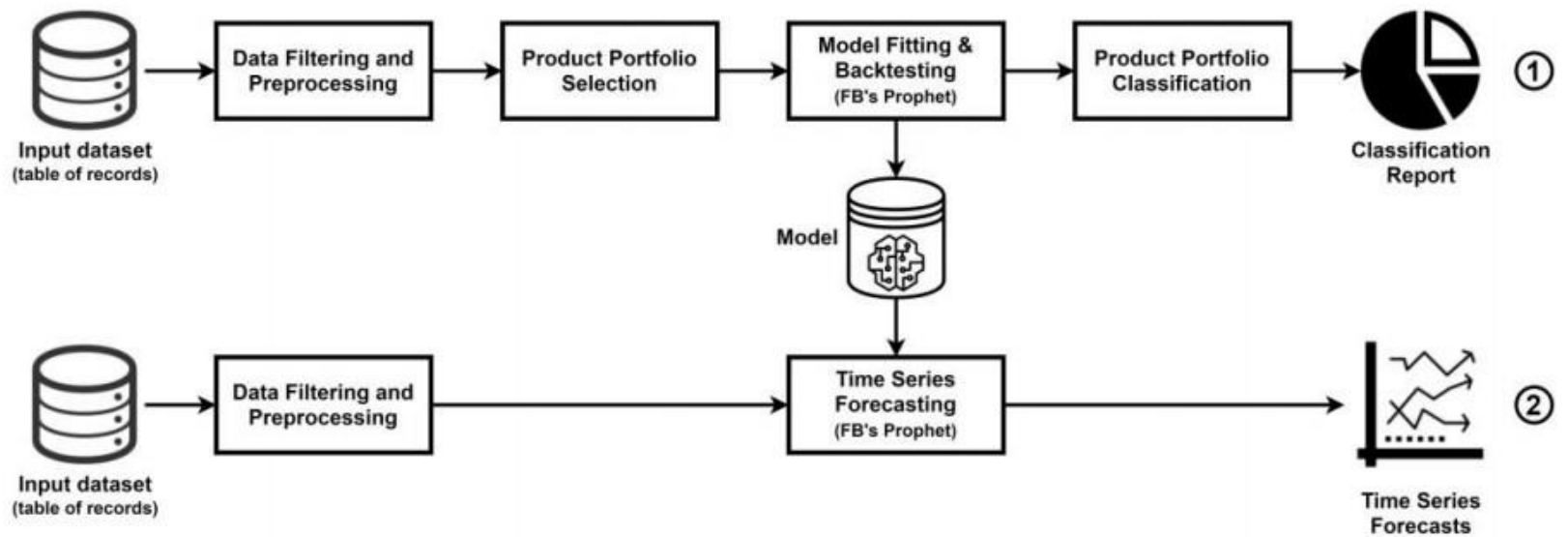
2)LSTM

Prophet

Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects.

It works best with time series that have strong seasonal effects and several seasons of historical data.

Lets see about the prophet in detail.



Input Dataset

- To develop a framework for sales forecasting the following columns were assumed to be available in the real-world input dataset which is structured as a table of records:
 - 1) item_code - unique identifier of the product in a portfolio
 - 2) date - date of transaction
 - 3) quantity - the quantity sold in a given transaction
 - 4) unit_price - the unit price at which the product was sold (optional, not used in forecasting)

	item_code	date	quantity	unit_price
0	5010010000001	2010-01-02	399	1.3300
1	5010010000001	2010-01-04	812	1.3380
2	5010010000001	2010-01-05	516	1.3310

The input to Prophet is always a dataframe with two columns: ds and y. The ds (datestamp) column should be of a format expected by Pandas, ideally YYYY-MM-DD for a date or YYYY-MM-DD HH:MM:SS for a timestamp. The y column must be numeric, and represents the measurement we wish to forecast.

Data Filtering and Preprocessing

- Several steps were taken during the preprocessing phase to transform a table of records into a convenient form:

filtering by date, in order to remove irrelevant historical data (e.g. it was decided not to

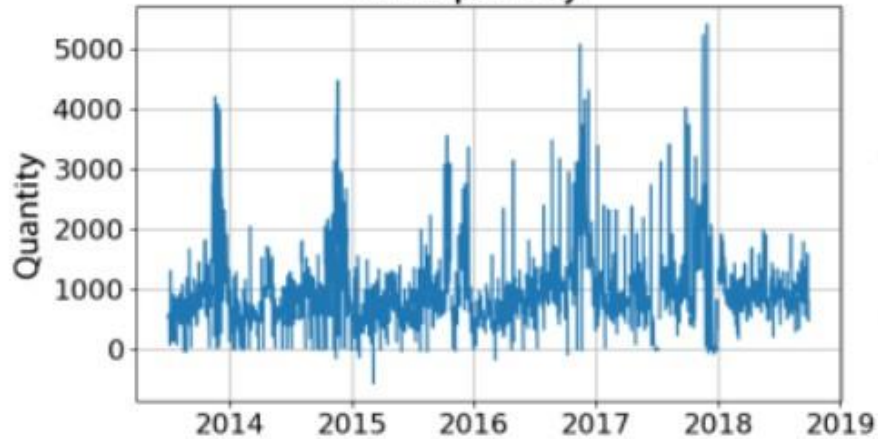
use more than six years of historical data)

conversion of quantities into the same unit (e.g. pieces, packs, bundles, pallets, etc.)

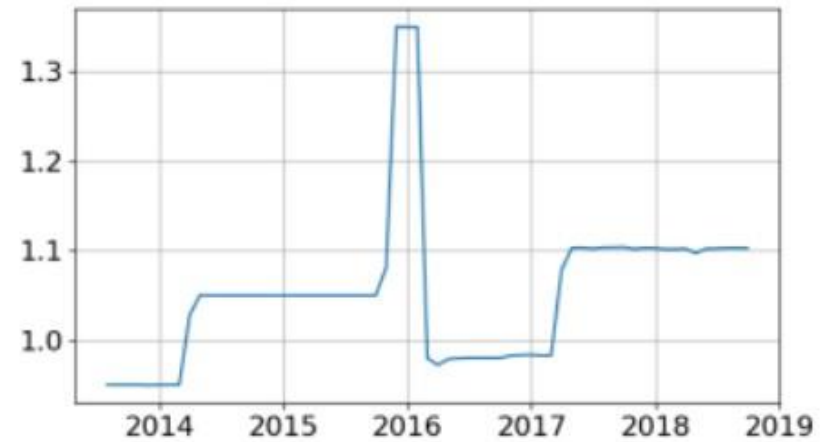
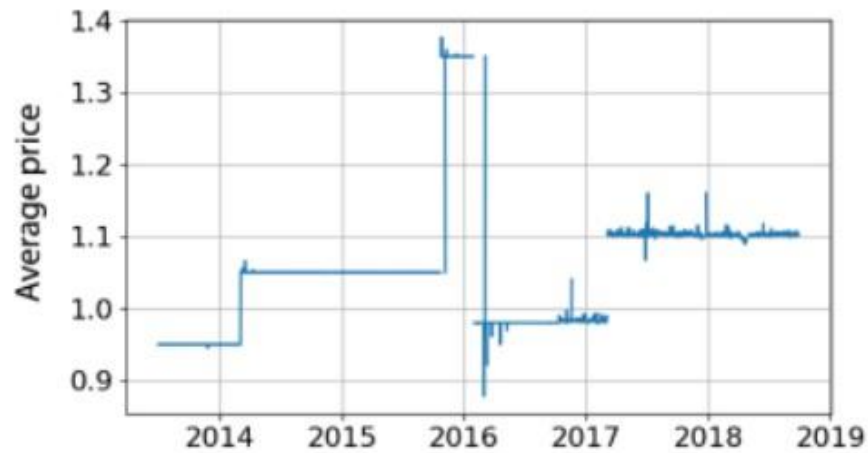
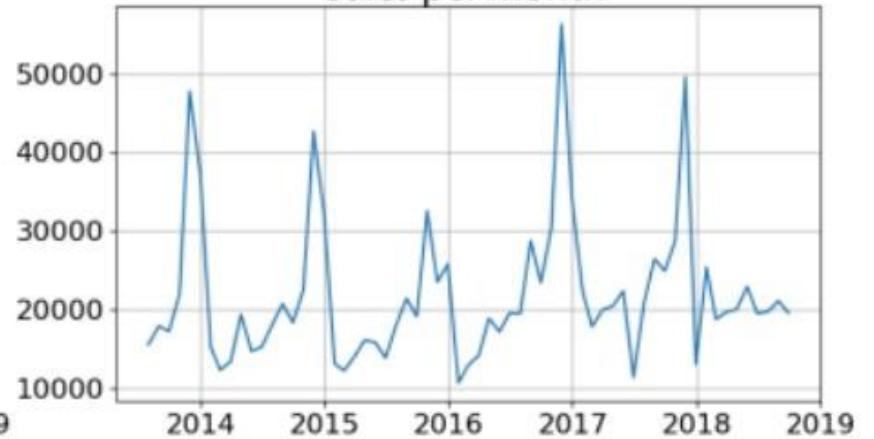
data aggregation in time domain at the product level (i.e. daily sales data were converted into monthly sales)

501009000004 | Ground black pepper, 50g (...)

Sales per day



Sales per month



Product Portfolio Selection

It was concluded that several different criteria can be used to sort the product portfolio by reference

- 1) the total profit per product over the year
- 2) the total financial turnover (i.e. net sales) per product over the last year
- 3) the total quantity sold per item over the last year

Performance Metrics

Two performance metrics are used for measuring forecasting accuracy, calculating the expected level of forecasting accuracy and classifying product portfolio accordingly:

- 1) the relative or percentage error (PE) for individual monthly/quarterly forecasts, and
- 2) the mean absolute percentage error (MAPE) for quantifying the overall accuracy.

The percentage error (PE), which can be calculated as:

$$PE = \frac{y_{forecast} - y_{true}}{y_{true}} \cdot 100\% \quad (1)$$

is mainly used to measure the accuracy of individual monthly/quarterly forecasting outputs generated by the model, while the mean absolute percentage error ($MAPE$), calculated as:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_{forecast}^i - y_{true}^i}{y_{true}^i} \right| \cdot 100\% \quad (2)$$