

CMPT 733 Big Data Science - Template for Capstone Project Idea

Automatic Hierarchical Time-Series Forecast at different Aggregation Levels for Fashion Products

Description

Sales or demand time series of a retailer could be organized along three dimensions: space, time, and product hierarchies. The spatial dimension captures the geographic distribution of the retail stores at different levels like country, province, city. The temporal dimension defines the chunks of time for which data is lumped together; for instance yearly, weekly or daily sales. And finally, the product hierarchy represents an administrative organization of products in some levels usually suggesting some degree of similarity: departments, categories, styles, etc.

In the context of a retail analytics software, the user might need forecasts at any of such spatio-tempo-product hierarchical aggregation levels; for instance city-monthly-department or store-weekly-style. The challenge, though, is that going up and down the aggregation levels, the characteristics of the time series (like its shape, patterns of seasonality, etc.) will change making it difficult to simply have a single time series forecast model. The idea of this project is to explore methodologies to cope with this challenge. There are at least two fronts:

- automatic choice of the appropriate model (either different types of models or different parameters for the same class of models) for an aggregation level
- create bottom-up or top-down consistent forecasts along each dimension
- find the optimal sweet spot of aggregation for more accurate forecasts

Datasets

The dataset for this project consists of 3-4 years of anonymized sales transaction data of a Canadian fashion retailer and will be provided by FIND. Data will contain product hierarchy information and lowest temporal resolution will be weekly.

Contact person

Hassan Saidinejad
Data Scientist
FIND Innovation Labs
hassan.s@myfind.ai



Please send the filled template to Jiannan Wang at jnwang@sfu.ca.