Sales prediction model

(Apparel Retailer)

- Integrated data from multiple sources into a data warehouse and automated the ETL processes for loading data from all sources on a regular basis
- Designed dashboards from the single and updated source enabling the executive team and departments to track critical KPIs related to Product Usage, Customer Success, Product Development, Sales & Marketing and Finance

Sales Prediction Model

ABOUT THE CLIENT

The client is a leading **apparel retailer** based in the United States

SITUATION

- Impromptu and unstructured promotion planning led to **low gross margins** and made it difficult to attribute the impact of promotional spending suggesting potential opportunity to improve the overall promotion strategy
- Merilytics partnered with the client to help **attribute the impact** of individual promotions and **predict the daily sales** based on the planned promotions.

VALUE ADDITION

- Evaluated and comprehended the **key factors** that will impact the daily sales such as prior promotional events, fashion season, proximity to major holidays and the type/depth of discounts.
- Estimated the historical impact of individual promotional events on daily sales to attribute the impact to each active promotion on a specific day
- Built **various algorithm** incorporating **forecasts from** a **LSTM and GRU neural network** into a **granular gradient boosting algorithm** to **calculate daily sales** for each product category.
- Customized the base algorithm to build specific models for each category for better accuracy

IMPACT

- The custom algorithm accurately **predicted daily sales at a category level** that helped the client to improve **inventory planning leading to reduced** Inventory costs and working capital
- This also enabled the client to run various **simulations to evaluate** the potential impact of a specific promotion, identify potentially weaker campaigns and improve promotion spend ROI

Approach / Methodology For Project Delivery

- Explore conventional and non-conventional **data sources** to identify potential features that will impact product sales
- Assess the relationship (distributions. T-test, P-test, ANOVA) between features and the target variable
- Reduce noise in the data, **impute missing data**.

- Use existing features as a basis to create **hybrid features** containing relevant information
- Compare the impact of hybrid features vs. basic features to gauge improvement
- Improve forecasts by exploring **auto-correlated features** as inputs to the model

- Build a **LSTM and GRU neural network** model to provide daily sales forecast. Optimized the forecast to minimize error and establish a range of forecast accuracy (i.e. the maximum length of forecast with acceptable accuracy)
- The forecast from the LSTM is used as a feature in a **CatBoost gradient boosting algorithm** along with all the other features to estimate daily product division sales.

- Provide a **summary of the impact** of promotions in the promotional calendar to the client
- Provide a toolbox that allows the client to run scenario analysis with **potential promotions** and estimate their profitability
- **Monitor model performance** and ensure error is within the acceptable range. Plan model refreshes to account for changes in market trends.

Key Variables Picked For The Sales Prediction Model

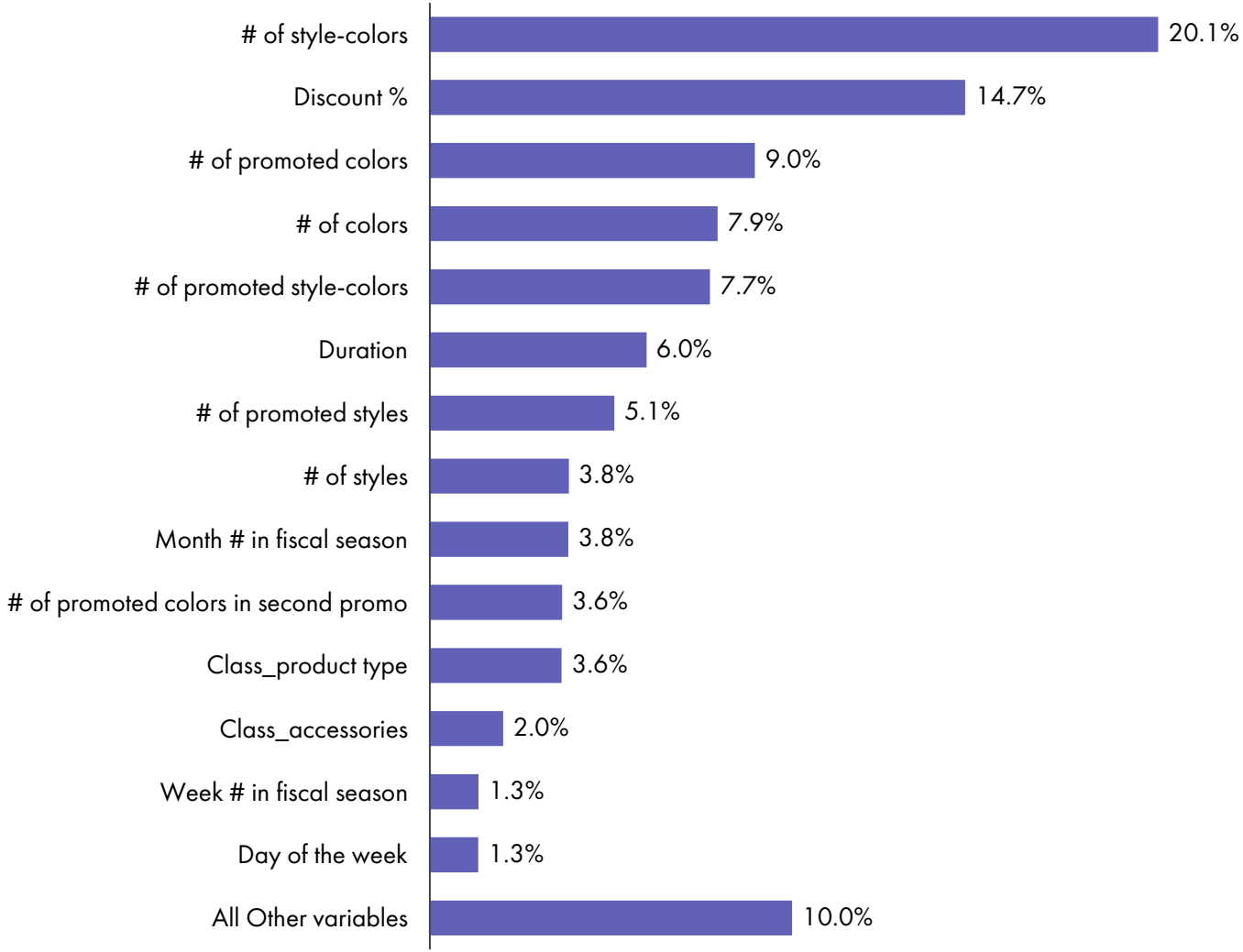
Dependent variable is total units' sales in a day (during promotion)

Category	Variable	Comments
Seasonality	Holidays	Impact of major holidays in the U.S. such as Black Friday, Christmas etc.
	Fiscal Season	Flag to identify FWH or SS season
	Week # in fiscal season	
	Month # in fiscal season	
	Day of the week	To differentiate impact on sales by day of the week
Product characteristics	Class	This is split into 4 variables to differentiate the target group (Girl, Boy, Adult, Toddler etc), category (sweater, swim) and product type (top, short, tight)
	Prior promotions	# of promotions on the same class in consideration in the prior 14 days
	# of styles	Total number of styles in the class under consideration
Promotion characteristics	# of promoted styles	Number of styles that are under promotion in the class
	# of promos in a day	Number of other promotions that are offered on the same day
	Promotion type	Flag to bucket promotions into Discount, Price point and BOGO
	Duration	Duration of promotion (1 day, 2 days etc.)
	Day of the promotion	Day since the promotion began
	MSRP	List price of the product in the store
	Discount %	Average % of discount offered in the class under consideration

Key Variables That Are Prominent In Predicting Daily Sales – Based On Catboost Model

ILLUSTRATIVE

Key variable importance[1], %



1. Denotes the relative importance of each variable used in the model to predict the unit sales per day

Predicted Daily Sales Vs. Actual Daily Sales

Actual unit sales vs. Predicted units for Category A, based on CatBoost algorithm

ILLUSTRATIVE