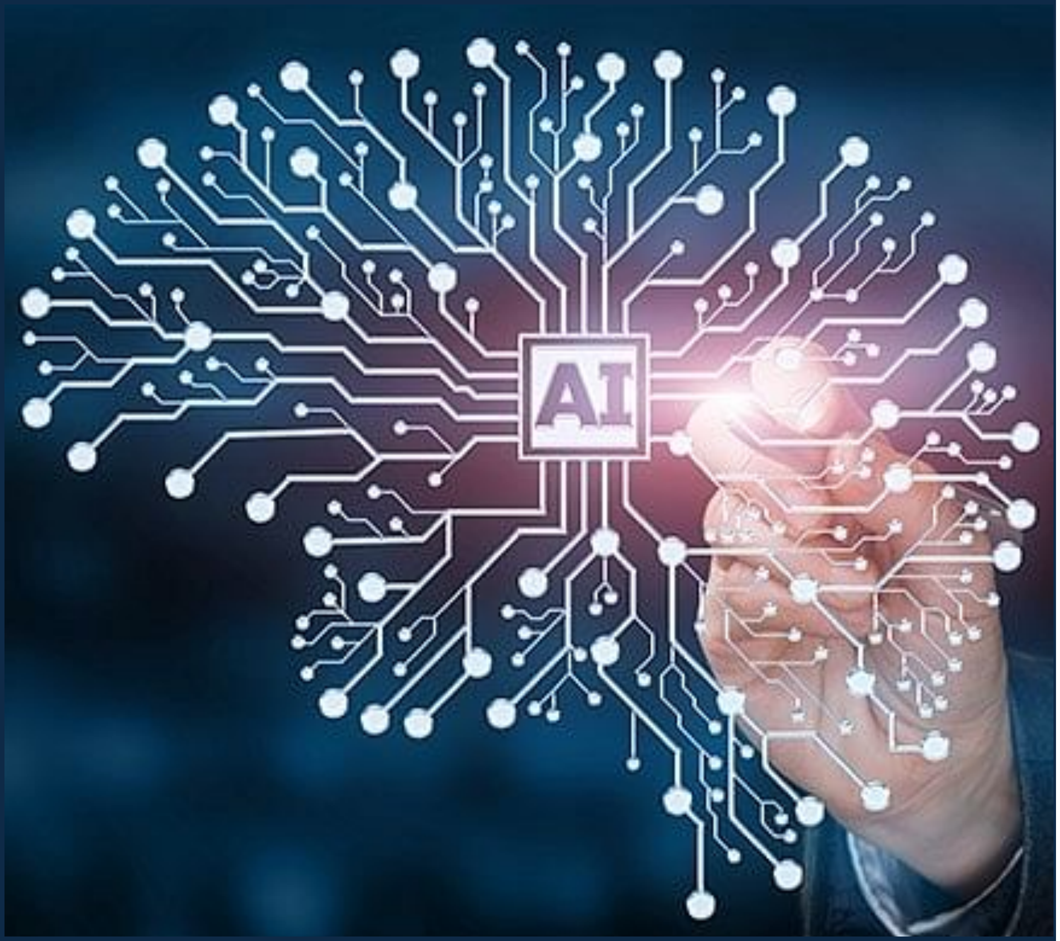
AI powered Assortment Planning Solution

Shankar Radhakrishnan

Mar, 2022



01 Context

02 Store Owners' Actions

03 Assortment Problem

04 Data Insights

05 AI Solution

06 Roadmap

Context…

- For a shop owner to maximize its revenue, there are few things to get it right
    - ensure that the products available on shelf though continuous replenishment
    - use the limited shelf space efficiently by placing the bestselling products
    - increase the purchase size by identifying up-sell and cross-sell possibilities though adjacent products
- At the same time, the real world is very dynamic
    - CPG organizations introduce new products all the time. While promotions are planned, how does one know which product could sell better?
    - How does one assort the right mix of products (new and established) to reduce inventory pressure of getting stuck with slow moving products?

Store Owner & Salesman' Actions

- Typically store owners have very scant information on these, and decisions are largely driven by human knowledge of trends and inputs from the sales representative. This is very prevalent in general trade market where point of sale information is not available.

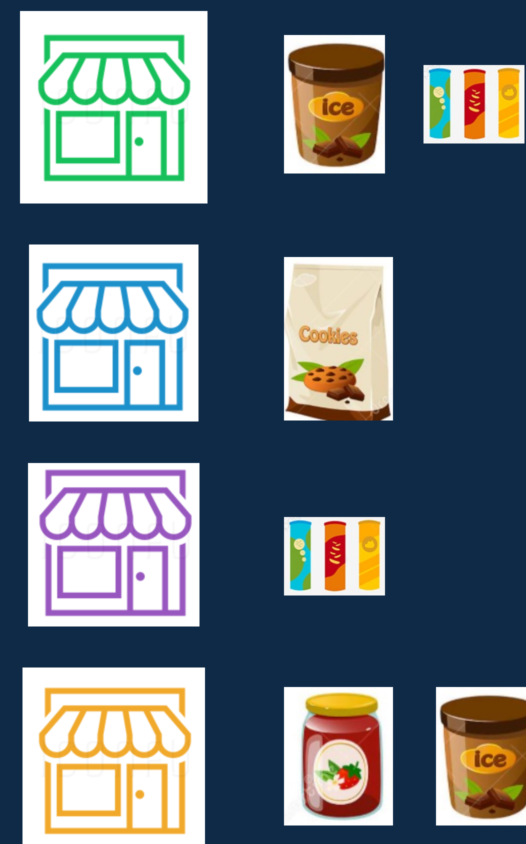
- Store owners must track and record these insights, conduct trials of assortment, determine appropriate price points and repeat such offers to convert to a effective sale

- Such actions involve lot of manual effort and insights out of such experiments have to be passed on to the brand's salesperson who will decide further on their brand' value

- The salesperson' action in such cases is very limited as they must correlate to SKU's that are on shelf, those in the inventory and predict the right quantity to replenish at the right time

- Such manual efforts are error prone and cannot scale as the product assortment grows and the demand patterns are diverse across store locations.

What is an assortment ?



Assortment growth and healthy on-shelf availability are critical to drive sustainable revenue growth

Most CPG companies build assortment but experience the leaky bucket problem where only less than 20% of their assortments tend to have a consistent repetition and depth

Salesman

Each Salesman visits at least 30 stores every day covering a distance of approx. 25 Miles each day

Salesman has approx. 5 to 7 mins time to spend in each store after taking into account of the travel time

Each Salesman has a hand held device which recommends the list of packs as a target to be sold in every store during the day

Assortment, Replenishment Decisions

This is primarily due to the salesperson who is given a large target to chase for each store while he has less than 15 minutes in most cases, to service a store considering the market time and the number of stores visits to cover every day

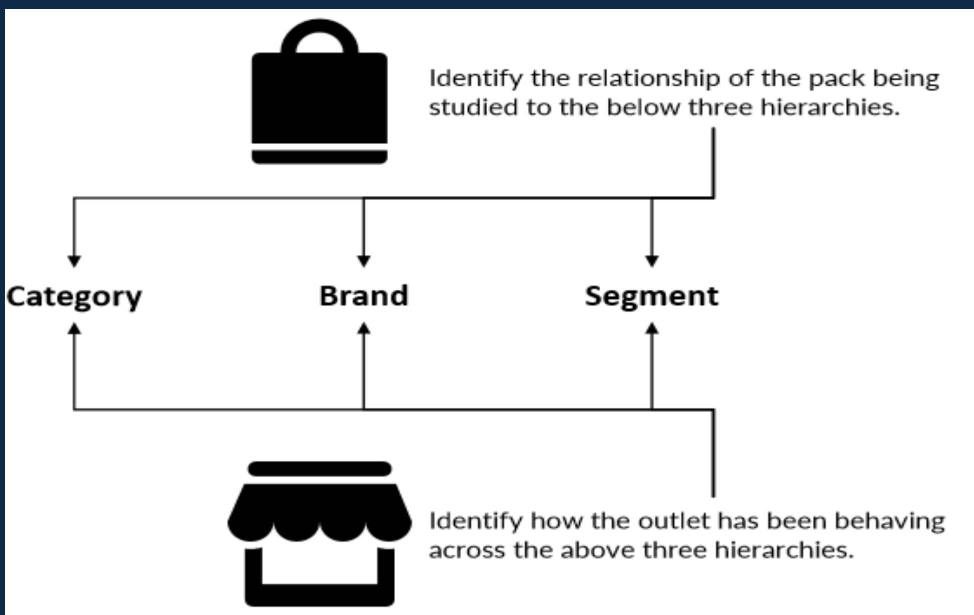
In this process the right assortment to be replenished on the shelf is missed thereby leading to loss of revenue and future assortment opportunities

Given that stores can cross source assortments from other channels, the above problem has other higher ramifications

It is therefore important to ensure consistency in service depth for each store, having the right assortment JTBD and identifying the right service window for each assortment.

Insights from data

Outlets react differently to different packs



Outlets chase different objectives across different pack types

Source of data – store fronts

Solution Overview…

Solution captures store and product behaviour and benchmarks store and product performance across the store segments in every local geography, at scale

Databricks’ distributed environment is used to compute and measure the store level behavioural attributes, product behavioural attributes and store neighbourhood behavioural attributes

The solution has embedded AI/ML algorithms to infer and develop various product and store relationships using which each assortment is prioritized for every store based on their last purchase behaviour

As part of the process, data from heterogeneous sources such as internal sales information and external information from third party sources are consolidated and synthesized

Solution Overview

Various KPIs using internal frameworks are computed to measure brand performance, category performance, premiumization, profitability and other aspects of store and neighbourhood attributes.

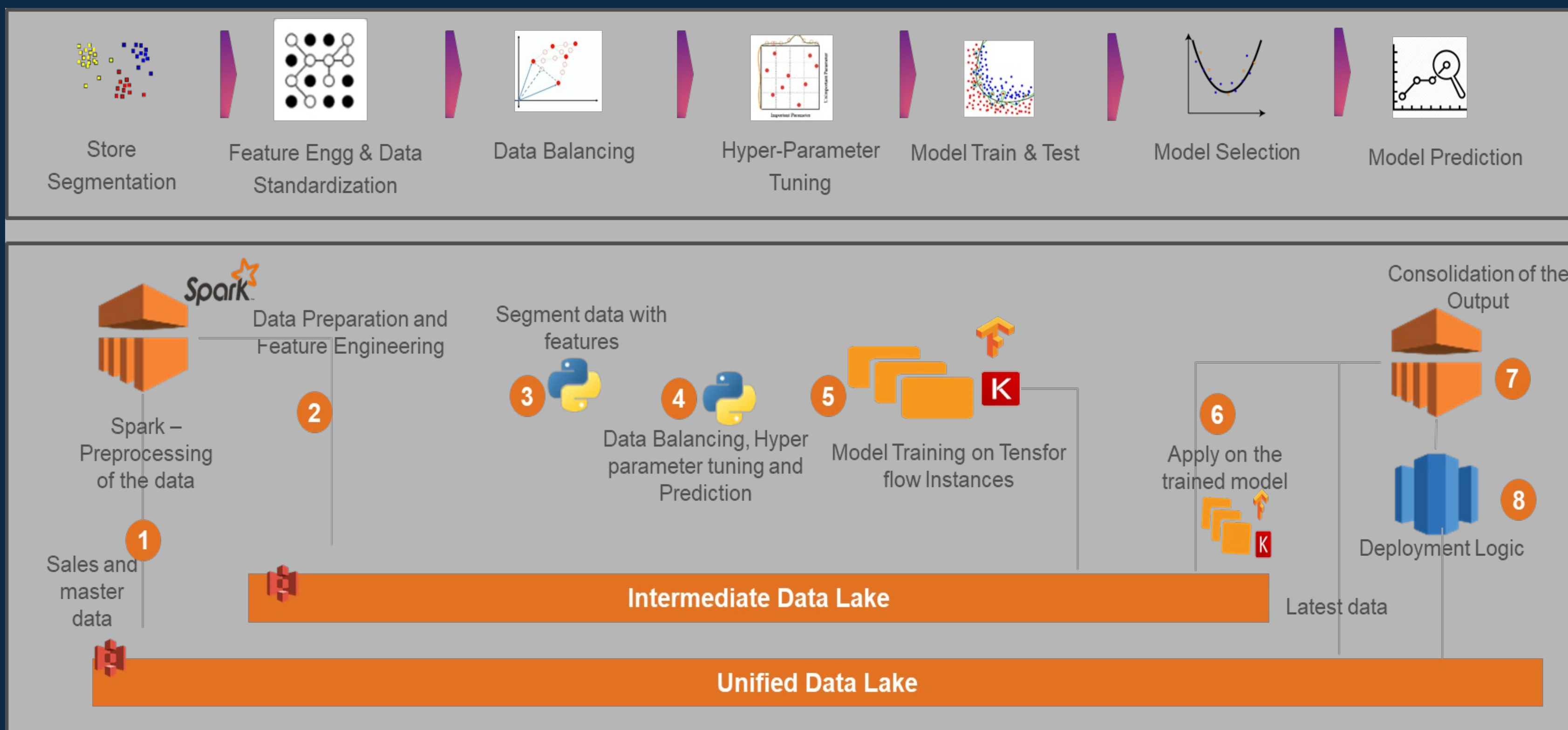
The data is then profiled and standardized before passing it to the recommendation model

Given that different data sets reside across multiple source systems, data pipelines are created to consolidate all the information in a common Data Lake

The solution can access this data from the Data Lake for the relevant behavioural KPI on Databricks. All the KPIs are computed using the distributed environment in Databricks for each of the Store profiles in a particular neighbourhood in a parallel manner. This led to a significant reduction in compute time and cost.

To ensure there is no class imbalance in the data, minority oversampling operations are performed. The core of this solution lies in the AI framework which has self-learning capabilities without being explicitly programmed while not over-fitting and under-fitting the data

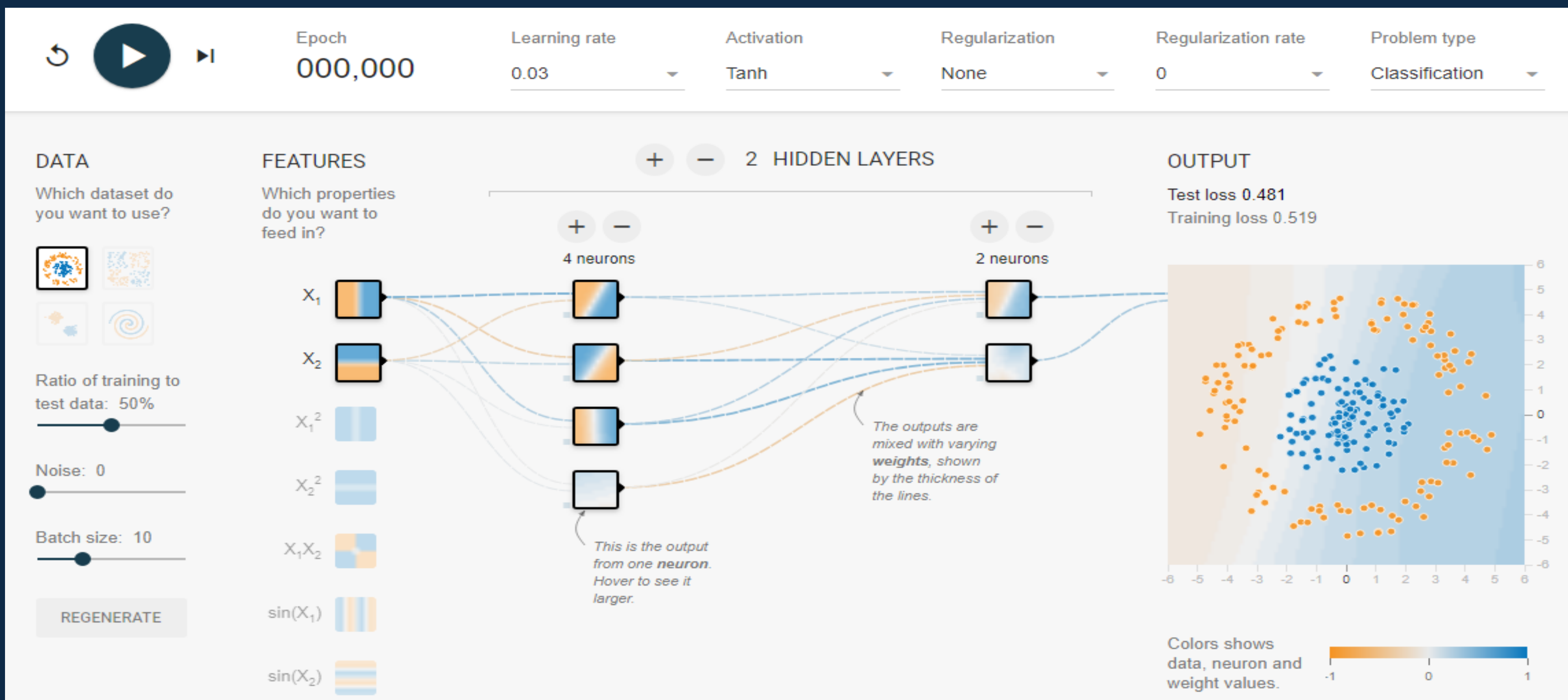
Solution Architecture



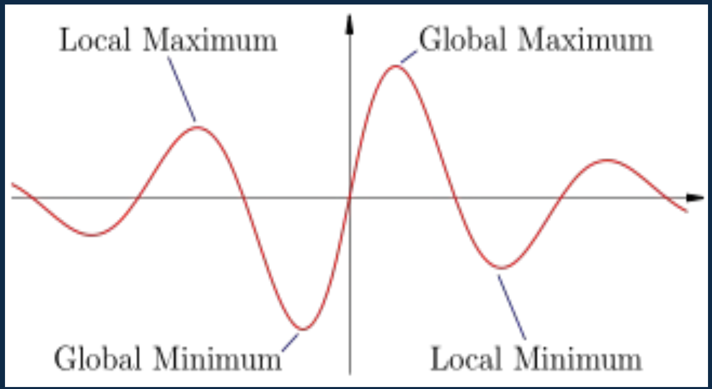
Solution Components

- Cloud platform
- Python, PySpark
- Databricks
- AI, ML, NN, DNN
- Open-source tools
- Data Lake on Cloud
- Reporting Tools
- Datasets (in-house and external)

Neural Net Components...



Neural Net Components



Learning Rate

Activation Function

Number of Neurons

Number of Hidden Layers

Batch Size

Regularization

- Relu activation function was used in the process given that from empirical understanding that they are less susceptible to vanishing gradients

- The hyper-parameters were optimized using a stochastic optimization approach to select the best combination of the hyper-parameters for the model execution process

- Approximately 900 Neural Network models were executed for all the store segments

Solution Highlights

- End consumer behaviour information - personalize store recommendations despite lack of end consumer sales information due to the nature of business trade
- Derive store behaviour from secondary level sales - understand store behaviour from sales data between the distributor and the store
- Build store segments dynamically - derive store segments dynamically by devising a framework to understand the number of clusters required under a geography based on data size and trade behaviour
- Understand opportunities using store attributes and pack attributes - derive features that drive store behaviour from the sales and the master information that is made available
- Leveraging AI based frameworks to predict using self learning algorithms - leverage AI based frameworks such as Deep Neural Networks to self learn on an ongoing basis with an automated program
- Leveraging parallelization using cloud - multiple Deep Neural Network model being executed for each store profile by harnessing the power of parallelization on cloud
- Leverage low-cost instances by building resilient orchestration framework that can handle interruptions - using low-cost spot instances
- Faster time to Insight
- Highly scalable and economic solution

Solution Roadmap

Salesman Route Optimization
Support more than 4000 distributors in enabling a scientific route design for their salesman

Pack Level Inventory Management for the distributors
Support more than 4000 distributors in estimating the upcoming month demand at pack level

New Product Launches
Support sales operations team to target new launches in specific stores to improve their launch hit rate

Salesman's Product Mix
Suggest the product mix of the salesman keeping in mind the product breadth the company caters to

Sales Estimation with Google Mobility Data
Estimation of sales using Google mobility data in order to counter challenges posed by lockdowns

Data Engineering and ML Engineering using Big Data on Cloud
Cloud based solution leveraging the on-demand and spot instance capabilities

Organizational Considerations

- This solution blends old world information collection, to new ways of tracking sales
- Store owner & Salesman has to feed data about stores, point-of-sale, products, promotions, inventory etc.
- The solution is location specific as the sale could wary between any stores
- For the organization that's implanting this solution, they should upskill the team on both AI, ML, Data Storage & Processing techniques and also CPG domain nuances
- A good training plan should be put in place and also, workshops with store owners and salesman will help the team to understand the dynamics from those on the field
- This solution stats away from tracking any personal information of the buyer, instead it focusses on the sale (what's been bought or not bought and there by stagnant)

Thank You