A Generalized Analytical Approach for Optimized Logistics Operations

Introduction: In today's world where online shopping has dominated the sales of almost all kinds of products. Given the flexibility and comfort to choose; its growing larger and larger. Online services are now a days reaching to the most of the part of our country. But given the limitation of supply chain management, we are still unable to do our best. For some part we are delaying more than usual since we don't have transportation hub nearby, the final delivery has to wait a lot before reaching the destination. Some part, we have more than sufficient services where we are unable to utilize them. It always happens when we lack of connectivity, where we have inefficient tracking of our progress, limitation with our standard check points and proper management. It is always been crucial have the best logistics facility for online shopping to be successful and earn and reach the customer's trust. So, we want to create strategies for both existing logistics and new potential logistics operation. We will focus on both improvement and implementation of logistics by these strategies. These strategies would be both efficient and profitable. This is a strategic model and is made of 3-sub models to create one unique idea. These sub-models are;

An Acquisition Model: This sub-model takes care of important data pattern extraction, Machine Learning Models and Market Strategies; it also takes care of future model (a model to predict certain behavior for new strategies).

Improvement and Implementation Model: This sub-model takes care of enhancement in logistics and operations; it also introduces new strategies which can optimize and utilize logistics operations. This model is not just limited to logistics operation but also takes cares of other dependencies. This model will also give us more data to improve first model as well as third model.

A Profit and Loss (PNL)/ Valuation Model: This is a validation model and a crucial one. Here profit and loss calculation will take place for each strategy that we come up with in first-model. Its results will decide which strategy to engage with. As each movement or change might require money investment so we calculate certain measures based on certain information which we will get in first sub model; this will give us good idea what to engage with and what to not.

Objective: This model works by inter-relating the information between these three models and comes up with best possible outcome. Each model works as a dependent as well as an independent unit to reach its central idea. Again the objective can be broken down into 3 sub-objectives/tasks.

Tasks of Sub-Model 1: Based on available data, this model has 3 sub-objectives. 1st collect all the available data related to the logistics services and operations. Data will have the information related to supply chain management, transportation hubs and sales information. The information will have Google map locations imbedded to it and as well as their financial information. So, this information will be used in both strategies as well as in valuation. The available information will help us in finding out existing strategies that we are currently engaged with. We will validate these strategies with our valuation model later. But mostly we will use this data to train our strategies. Any new strategy will also add information to our model. 2nd, we will frame this data into our standard format. Since, there will be some required information which we will need to have; so, data completion will be important part. 3rd, we will check data sanity, which will separate confidential and public data and process them differently.

Tasks of Sub-Model 2: This model will help in implementing strategies; strategies that will be based on collect data in 1st sub model. The strategies will be both new and old. Old strategies mean, pattern extraction and decision analysis of available past data. And, how's these strategies working out for us will be decided by our valuation model. Enhancement on old strategies will be decided by their results from valuation model. Valuation model will provide us a lot of measures to decide (will explain in third sub model) whether should we retire current settings falls under it or need improvement.

New strategies will also depend on data pattern we have; and similar validation will be done for new strategies also. Some of the strategy will also have real time data dependencies and their validation will be practical in impact.

Tasks of Sub-Model 3: This model is responsible for justifying our actions aka strategies. For each strategy, we will calculate the profits/losses, sanity measures, cost measures and other process related variables. In order to evaluate PNL, we will create some valuation measures, by the use of appropriate use of these measures; we will get the final predicted outcome of strategy. These measures are both of logistics service side as well as customer's side. How these measures are implemented will be explained in steps involved section. Stating the obvious, the best strategies we come across, we will do virtual scenario regression before going for real life implementation. This will be done by creating virtual scenario for which we are creating our strategies and run our regression recursively until we have our optimized solution.

Step Involved: Again for the ease of explanation, we can break this into three part or should I say, implementation of 3-sub modals. Dependency between models will be explained in both implementation and flow chart.

Model-1 Implementation: This step includes pure data mining tasks and heavy use of machine learning and AI and Google maps APIs.

1. Data related to logistics supply chain management. As mentioned above data will be collected stating from the online purchase till delivering the product. The information we will collect will have, Google map information of each point, where product is being forwarded to different locations, warehouses and location of delivery point. Sales information and whether we also other information related to product being delivered (Top level categorization not detailed information). We will also include the information related to warehouse.

2. 15+ machine learning and AI models will be used to train 15+ models, for real time processing one of best performing model will be used, and rest of the model will be used in background for multi-class classification. Selection parameter may include but not limited to- coverage, accuracy, performance, market value and support value for profit and loss etc. These models will help in patternistic strategies. Which are already there and their valuation will give brief overview of it.

3. We will use clustering method for few strategies for which we collected the Google map data, this clustering method will be more crucial in few optimized strategies.

4. Now, combining all the data we have, we will complete all required data, missed information can be completed from old and reference and records. We will map down our information in our standard pattern and also will keep public and confidential information separately although we might use combined information to make strategies or to extract our valuation measures.

5. We will use Python Libraries to do all our work here, from collection till data preprocessing and mapping.

Model-2 Implementation: This model is more focused on strategies side since it works mostly on available data from 1st model.

1. First this model create feature matrix corresponding to products (product that brand sells) from the collected data. The data mapped in the form of feature matrix will be used for clustering, and based on the data cluster and their behavior we will consider the strategies.

2. These strategies will be validating by valuation model. Since the initial size of the cluster is going to be large, we will move from category of products to specific products, location to specific location from warehouse to specific deliver services and product management. This will reduce the errors in feature matrix. Once we done with the validation of recommendations, we will transfer this to above model.

3. New strategies will be formed by use of model-1 data, which will be twisted with virtual scenario to consider all possible situation and outcome.

4. This final step includes a background process, this process will keep a trained model which will be based on type of deliveries we are doing, like shipping or drop shipping; since, delivery can be from the vendor itself or the third party both in case has different performance scenarios.

5. Type of strategies will include like location strategies for warehouse, where sales, deliveries and warehouse Google location and product details will be key points to consider.

6. Strategies related to staff and management will include more information related to warehouse and their upstream and downstream supply data, which will be used in management strategies. And so on.
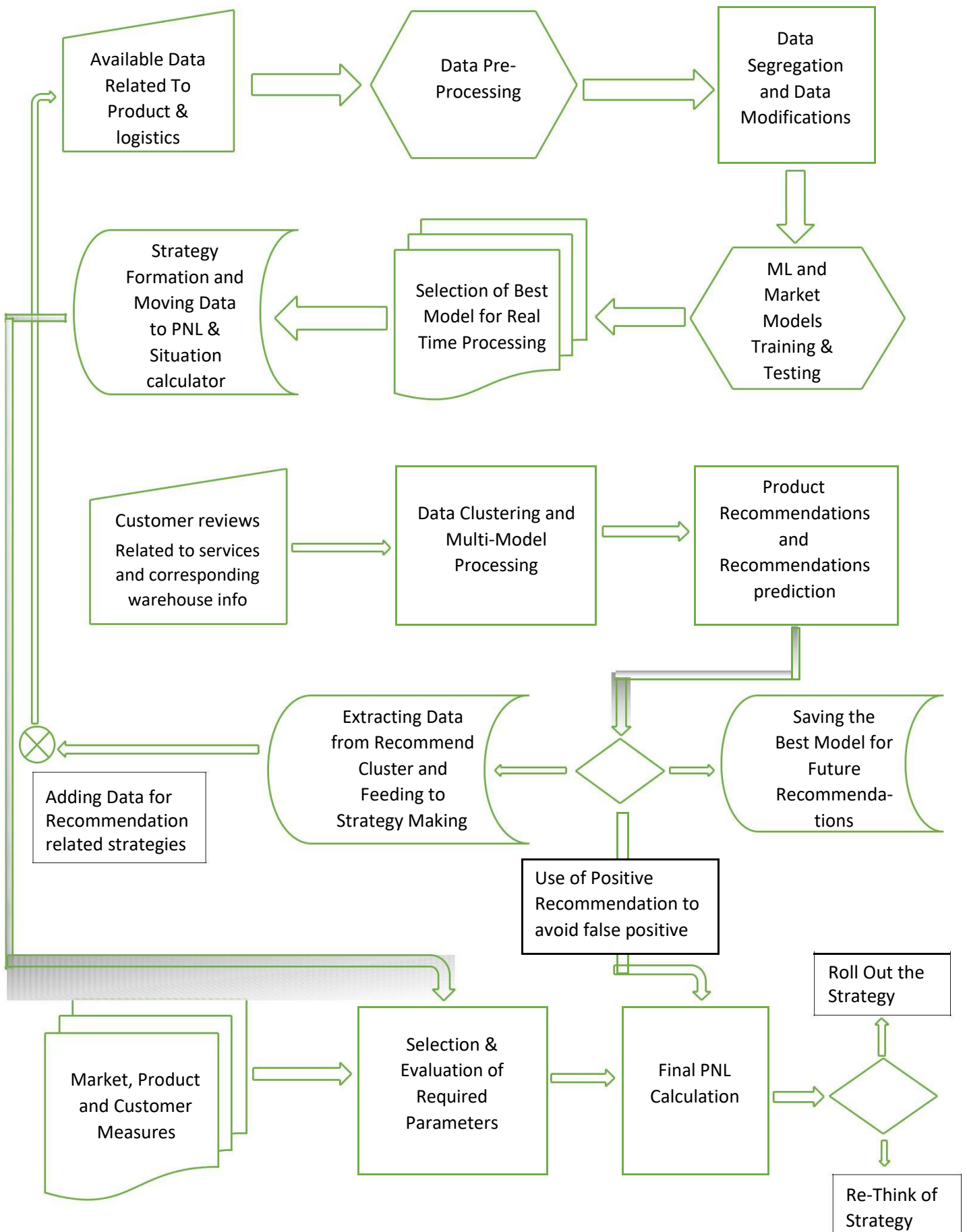
7. Few of the strategies will come down with the help of data pattern. As data changes our strategies will change too and this all will be automatic process. We will only see results on UI which certain necessary fields.

Model-3 Implementation: This model is our validator, a model to calculate profits/losses from our strategies. It is not necessary that we will only engage with profitable strategies, sometimes lossy strategies might help in future benefits. So, we will keep all the strategies and their PNL reports and let the brand decide itself which strategy to engage with.

1. The PNL calculation needs product, logistics facility and supply chain management related measures that will be calculated from corresponding data. These measures are mathematical implementations of data patterns, such as price behavior, quality behavior, customer rating curve, product classification, service etc. These measures will be standard but the data they will be extracted from might change based on strategy.
2. We will train multiple measures for better outcome and accurate valuations. At least 10 related to each filed which we are depending for valuation.
3. For the prediction of our future strategies we will also implement measures for future scenarios.

Most of the process will be automated; every strategy will have their PNL. Decision making can be automated or can be left manual; it will be decided by brand and type of strategy. This entire process will be integrated on a UI which will reflect important results we need to consider after the valuation of any strategy.

Available Data Related To Product & logistics

Data Pre-Processing

Data Segregation and Data Modifications

Strategy Formation and Moving Data to PNL & Situation calculator

Selection of Best Model for Real Time Processing

ML and Market Models Training & Testing

Customer reviews Related to services and corresponding warehouse info

Data Clustering and Multi-Model Processing

Product Recommendations and Recommendations prediction

Extracting Data from Recommend Cluster and Feeding to Strategy Making

Saving the Best Model for Future Recommenda-tions

Adding Data for Recommendation related strategies

Use of Positive Recommendation to avoid false positive

Market, Product and Customer Measures

Selection & Evaluation of Required Parameters

Final PNL Calculation

Roll Out the Strategy

Re-Think of Strategy

Technology Used:

Mostly, we will work with Python, R and their extensive libraries. SQL, XML and HTML and JAVA Script will be used .

Benefits:

1. Use of multiple ML and AI models introduces less error, high accuracy and effectiveness of the model. Since best performing algorithm plays the role in real time processing, any new best performing algorithm can be introduced in main frame for real time processing and rest algorithms/models can be used in background for optimization and process enhancement.
2. Since its main frame uses python and R which is easier to modify and new changes can be rolled out easily. This model can for any kind of strategy as long as data can be mold into standard format.
3. It has ease of development, it has less maintenance cost and it can be fully automated (in full model implementation). Division of the process into part introduces the flexi option of adding some other effective part. Data limitation are also resolved by introducing second sub model.
4. Keeping dependency between models for less error is kind of advantageous since it will rule out false positive strategies. Selection criteria are being used to make sure we can make best out of it.
5. PNL plays important roles in figuring out the best strategies, and these trained model can be used in future trends and strategies and hence immune to data patterns changes.
6. These sub-models can also be extended into fully functional separate models for different purpose like a different system for customer recommendation so lot of possibilities.

Extra:

Please have reference of some of our previous project.

Data Mining Project, Future Trends Prediction case study, Sentimental Analysis (Previously shortlisted idea).

Team Data Geeks

Maneesh Kumar Meena