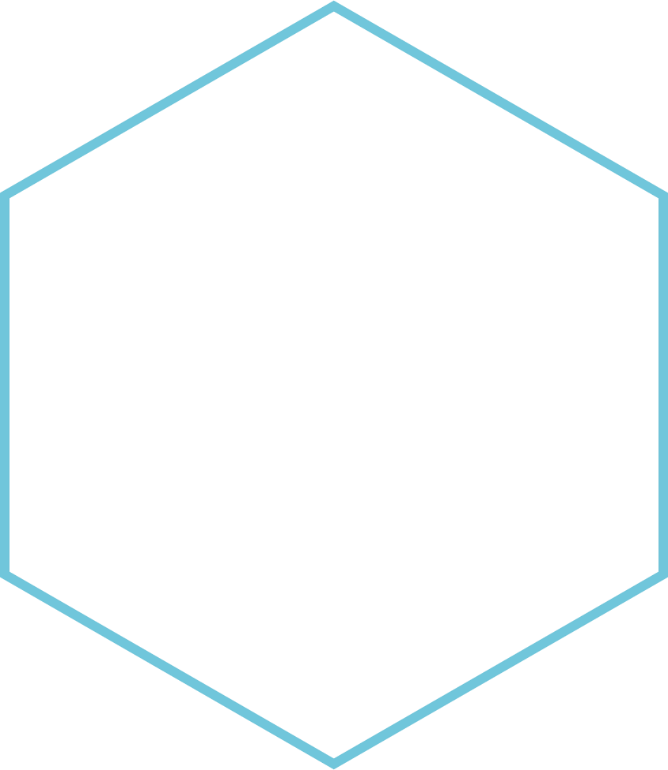


|  |
| --- |
| **Recommendation Engine for eCommerce** |
| Big Data Analytics – Batch 3 |
| *This project is meant for developing a recommendation engine to predict products likely to be purchased by e-commerce customers.* |
|  |





|  |  |  |
| --- | --- | --- |
| eCommerce Recommendation Engine |  | CRISP-DM • • • |
| CRISP-DM is an open cross industry standard process for data mining and provides a structured approach for planning and execution of data mining projects.The CRISP-DM project lifecycle is illustrated below for reference. Image result for crisp-dm We are choosing CRISP-DM versus ASUS-DM as ongoing optimization is not in scope for this exercise. |
| Project OverviewPProject OVerview |
| Context XYZ is a retail giant with operations with 250+ stores in 120 cities in India. It has presence in eCommerce space as well. It sells products across different categories like Electronics, Movies and music, Home and furniture, Home improvement, Clothing, Footwear, Jewellery, Toys, Health and beauty, Pet supplies, Sporting goods and fitness, Craft supplies, Party supplies, Grocery etc.  XYZ is looking for cross-sell / up-sell opportunities to existing and potential new customers and would like recommendations on top 20 products likely to be purchased by their customers. To answer these questions, we are developing a recommendation engine to help XYZ predict top items likely to be purchased at a customer level. The business value expected from this project is to increase XYZ’s chances of selling more products. This can be quantified as the opportunity to sell the top items predicted for likely sale by each customer and the cross-sell / up-sell opportunity for XYZ’s e-commerce customer base will be derived as part of this project. Approach CRISP-DM methodology will be used as a framework to execute this project. This consists of the following steps:  **Business understanding**: Understanding the need and drivers for recommender system - whether this is being done to turnaround declining revenues, is there a focus on a specific set of customers or products etc.  **Data understanding**: Exploratory data analysis and feature extraction  **Data preparation**: Data cleansing and feature engineering  **Modeling**: Feature Selection, developing product mix / recommendation models to suggest cross sell and up sell opportunities by customer.  **Evaluation**: This will be an iterative activity with the modeling effort. We can also evaluate this with the client if possible.  **Deployment**: The models will be developed and deployed using infrastructure in IIMB DCAL labs before handing over to the client. Data The data related to 7 stores in Amritsar, Hubli, Indore, Jamshedpur, Luthiana and Madurai. 8million rows of product data and 100’s of 1000’s of transactions data is already available. Scope The recommender system will be built based on the purchasing patterns of customers in India based on the datasets provided and any other publicly available and relevant data which the project team finds useful for this purpose. |