Low Level Design (LLD)

Analysing Google Apps Store dataset

Presented by:
Anwesha Das

Document Version Control

Date Issued	Version	Description	Author
10th July 2022	1.0	First Version of Complete HLD	Anwesha Das

Content:

1 Introduction

1.1 Why this High-Level Design Document?

The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the House Price Prediction dashboard. LDD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

1.2 Scope

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organisation may be defined during requirement analysis and then refined during data design work.

1.3 Project Introduction

In today's scenario we can see that mobile apps play an important role in any individual's life. It has been seen that the development of mobile application advertising has an incredible effect on advanced innovation. Having said that, with the consistently developing versatile application showcase there is additionally an eminent ascent of portable application designers inevitably bringing about high as can be income by the worldwide portable application industry. Developers and users play key roles in determining the impact that market interactions have on future technology. However, the lack of a clear understanding of the inner working and dynamic of popular app markets impacts both the developers and users. With enormous challenges from everywhere throughout the globe, it is basic for a designer to realise that he/ she is continuing in the right heading. The Google Play Store is observed to be the biggest application platform. It has been seen that in spite of the fact that it creates more than two-fold the downloads than the Apple App Store yet makes just a large portion of the cash contrasted with the App Store. In this way, I've scratched information from the Play Store to direct my examination on it. Actionable insights can be drawn for developers to work on and capture the Android market

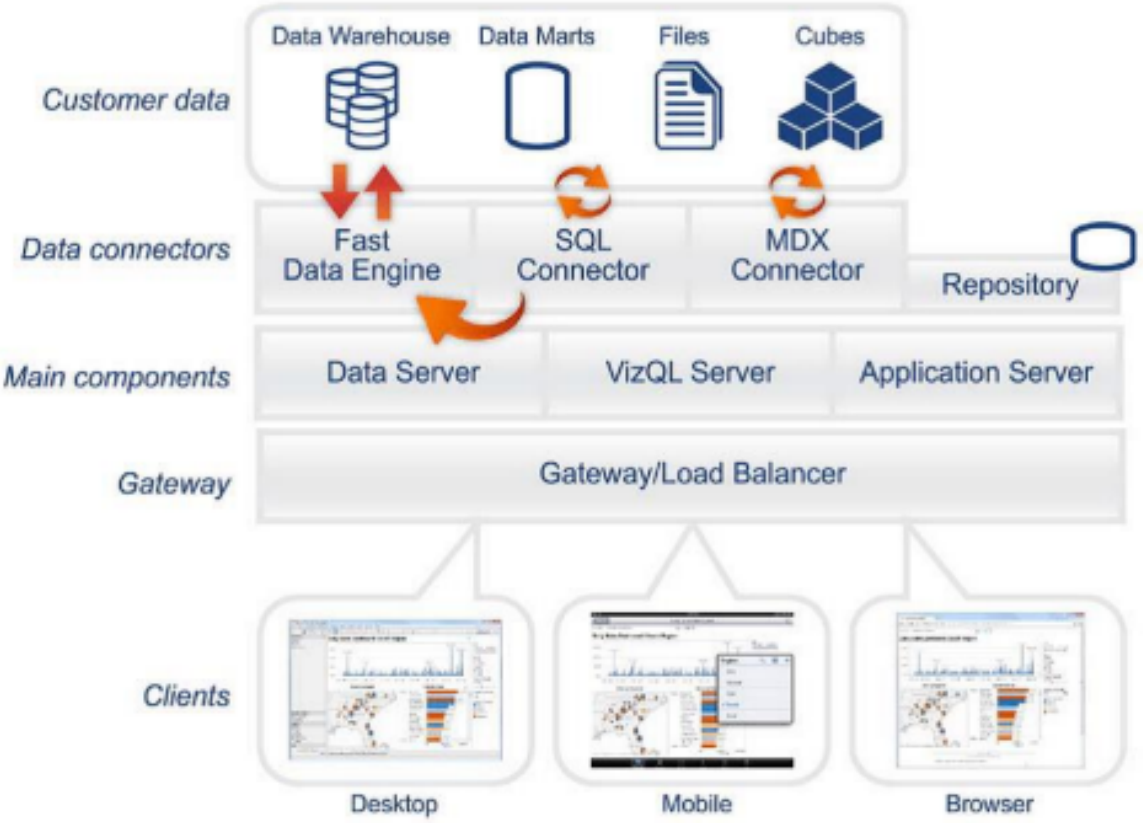
2 Problem Statement

Technology is an increasing need nowadays and used everywhere. One of the features of Technology is android. Which we all use in our daily life. Android is a mobile operating system based on a modified version of the Linux kernel and other open source software, designed primarily for touchscreen mobile devices such as smartphones and tablets. It is difficult to predict the growth of an app and know the state of the industry. There should be a way which will help app developers predict their number of installs or investors who want to pick out the next big app. Companies may run beta focus groups, or app developers may receive feedback from testers and get certain amounts of reviews. We will use this and some knowledge about the app to predict its success. Knowing the number of installs can help developers and business managers because they can predict the profit. This project's result may show the importance of reviews to apps in the market as it could be one of the determining factors for the number of installs.

3 Data Information

We have two datasets from Kaggle for app reviews; one is the list of apps with information. It has information such as app name, category, rating, and more. And the other is a list of reviews for each app with the sentiment if that particular content of the review was positive, neutral, or negative. Unfortunately, we could not directly use these two files as they are not joined.

4 Architecture



4.1 Working

Raw Data Collection:

The Dataset was taken from iNeuron's Provided Project Description Document.
https://drive.google.com/drive/folders/165Pjmfb9W9PGy0rZjHEA22LW0Lt3Y-Q8

2. Data Pre-Processing:

Before building any model, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict. Model performance depends on the

quality of data feeded to the model to train. This Process includes:

a) Handling Null/Missing Values

b) Handling Skewed Data

c) Outliers Detection and Removal

3. Data Cleaning:

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.

a) Remove duplicate or irrelevant observations

b) Filter unwanted outliers

c) Renaming required attributes

4. Exploratory Data Analysis (EDA)

Exploratory Data Analysis refers to the critical process of performing initial investigations on data to discover patterns, spot anomalies, test hypotheses and to check assumptions with the help of summary statistics and graphical representations.

5. Reporting

It is a most important and underrated skill of a data analytics field. Because being

a Data Analyst should be good in easy and self-explanatory reports because our model will be used by many stakeholders who are not from technical background.
a) High Level Design Document (HLD)

b) Low Level Design Document (LLD)

c) Architecture

d) Wireframe

e) Detailed Project Report

f) Powerpoint Presentation

6. Modelling

Data Modelling is the process of analysing the data objects and their relationship to the other objects. It is used to analyse the data requirements that are required for the business processes. The data models are created for the data to be stored in a database. The Data Model's main focus is on what data is needed and how we have to organise data rather than what operations we have to perform.