CarSpy

Wang Changqin, Zhang Haolin, Zhang Lei, Liu Haoyang
CS5224 Cloud Computing
AY2021/22 Semester 2
Department of Computer Science
National University of Singapore

An Intelligent Recommendation Platform to find your Dream Used Car in Singapore

Motivation

In Singapore, the necessity of reducing air pollution and maintaining smooth public traffic is highly valued. As a result, purchasing a brand-new car can be very expensive due to policy constraints. As an alternative, used cars can not only meet people's demand of owning their own cars but also fit their financial budget.

Innovation

Every recommendation made does not only consider the filter conditions that the user set up on the search page, but also analyze what kinds of car posts the user viewed frequently and even starred to the favorite list. This makes our service become a more user-friendly intelligent recommendation system, instead of a simple search engine.

Objective

We aim to build a cloud-based intelligent system that can provide users with customized recommendations on u-sed-car postings. We hope by using our service, users can find satisfactory used cars with fair market price and matched features.

Target User

Our web application targets the young population, taxi drivers, and low-income families. Used cars always fit these groups for a rather low price. As users of our web application, they will no longer be frightened by exorbitant prices or worried about not knowing the used car market. CarSpy could be a perfect choice to tackle their budget problem.

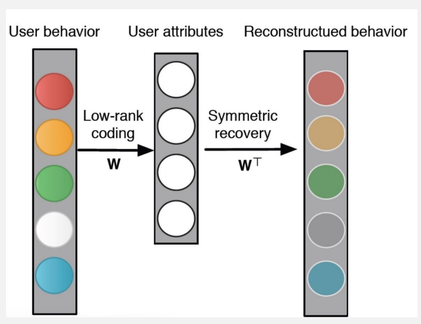
Data Acquisition

Original Dataset

Our original Singapore used car data is obtained from the Kaggle dataset: 100,000 Singapore Used Car Dataset, which contains the scraped data of 100,000 used cars listings, which have been separated into files corresponding to each car manufacturer. This dataset has 10 variables, including model, price, transmission, and mileage, that can be used for our recommendation algorithms.

Scheduled Scraper

We schedule a daily scraper for the used car data on CARRO to obtain daily data regarding Singapore's used cars. The daily data and the raw dataset files are processed by a CloudFormation function to store in DynamoDB.



Recommendation

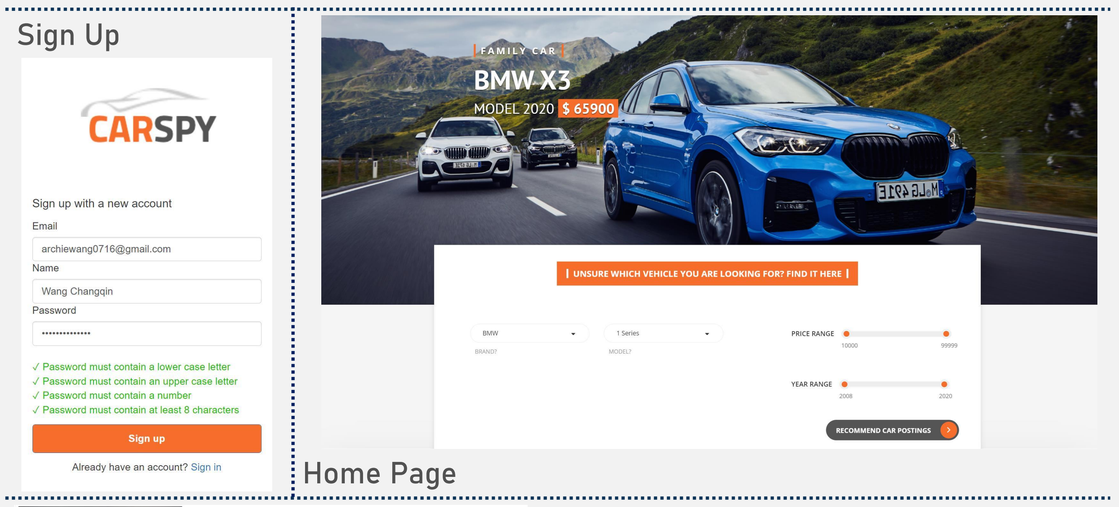
Linear Low-rank Auto-Encoder

CarSpy uses the Linear Low-rank Auto-Encoder (LLAE) algorithm to extract the hidden features of each user. Meanwhile, we introduce the low-rank constraint term, to avoid the undesired correlations between different users.

Cold Start

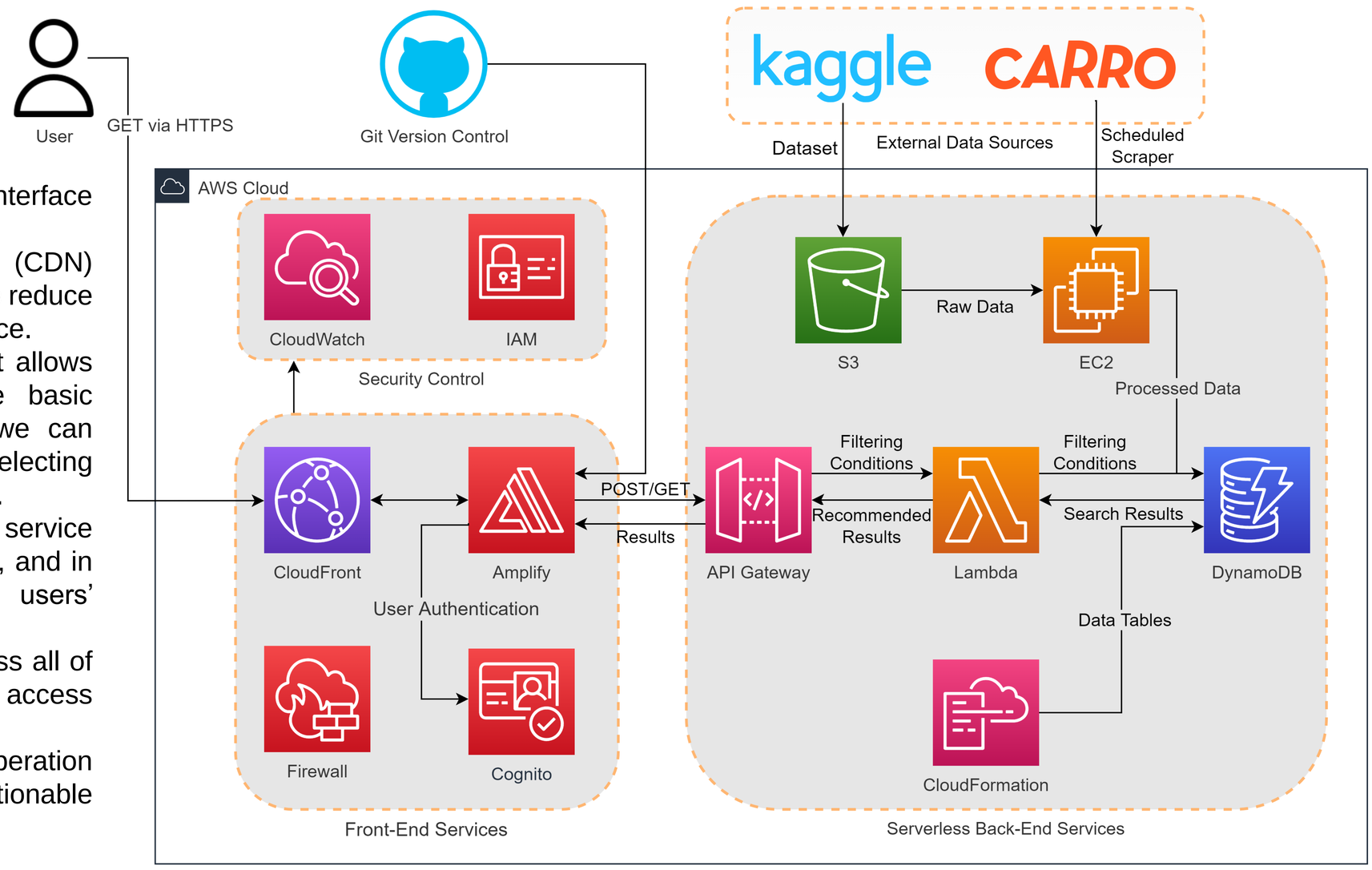
CarSpy leverages the user cold-start method for our recommender system when there are few interactions for the newly registered user. We use Content-based Filtering, extracting the content of each car item by PCA, and then the Euclidean distance is calculated for the recommendation.

User Experience



Implementation

AWS Amplify with GitHub hosts a full-stack web application by linking it to our GitHub repository and deploying our front-end code files to it.
AWS API Gateway provides a bi-directional interface connecting the front-end logic to the back-end logic.
AWS CloudFront is a content delivery network (CDN) service that deployed our Amplify app which can help reduce the latency and improve the security of our web service.
AWS Lambda is a serverless compute service that allows us to create Python functions. It provides the basic packages like *NumPy* and *Pandas*, with which we can implement our recommendation algorithms after selecting data from AWS DynamoDB using the *Boto3* package.
AWS DynamoDB is a fast, flexible NoSQL database service for single-digit millisecond performance at any scale, and in CarSpy development, we use it to store our users' information and both our cars' detailed information.
AWS IAM provides fine-grained access control across all of AWS, we create several roles for developers to access multiple kinds of services we have used.
AWS CloudWatch monitors and observes the operation situation of CarSpy and provides us with data and actionable insights to monitor our application.



Revenue Models

Subscription Based

Subscribed users can enjoy the premium recommendations.

Subscription Price:
Monthly: $19.9 ($19.90/month)
Semi-Annual: $79.9 ($13.32/month)
Annual: $119.9 ($9.99/month)

Advertisement

CarSpy provides advertisement services for used car trading platforms and individual sellers.

The Ads will be recommended to target users with a higher priority if the users search for similar items.

Cost Breakdown (Monthly Estimated)

Amplify	API Gateway	CloudFormation	Cognito	DynamoDB	EC2	Lambda	S3
Data served per month (15 GB)	Requests (500,000 per month)	Total number of operations (2,000 per month)	Number of active users (5,000 per month)	On-Demand, Data storage size (10 GB)	Instance type (t4g.large), Storage amount (30 GB)	Number of requests (500,000 per month)	Standard storage (5 GB per month)
$12.37	$0.5	$0.9	$6.75	$6.25	$33.73	$0.11	$0.13
$60.74							