

Name,

Instructor's Name,

Course Name

Date

Cloud Data Management and Recommendation Systems

Abstract

As data generation continues to grow, companies in public and private sectors are taking advantage of cloud computing technologies to manage and maintain data efficiently. The intrinsic value of a business today is the ability to have data and make insightful information from it. Due to its highly scalable resource, Cloud computing is deemed to host and manage recommendation systems. Through recommender systems, businesses can efficiently and effectively manage data.

Keywords: Data management, Recommender System, Hadoop, File Systems

Introduction

As data generation continues to grow, companies in public and private sectors are taking advantage of cloud computing technologies to manage and maintain data efficiently. Intrinsically, companies are today competing based on how much data they have in storage and the insightful nature of that data. Data origins not only from customers and products but also from the supply chain, competitors, or any other data that can fall under the big data umbrella (Çano, np). Cloud computing is currently considered as one of the best technology to store and manage data of this nature.

The data growth rate in cloud computing has led to the realization of recommendation systems to help and identify different aspects of cloud computing. This is also due to the need to filter and find data addressing a given domain. Recommendation systems have been developed and designed to address various business needs. In particular, recommendation systems are mostly used in academia

and banking institutions. The goal of this research paper is to discuss how data is managed in the cloud and various applications of cloud-based recommendation systems.

Cloud Data Management

Data management is an administrative process that aid in storage, retrieval, validation, and protection of data. Data management is a necessary process in any organization as it allows businesses to stay organized and run smoothly. In a cloud environment, there are three computing models under which a company can decide to host their data. Software as a Service (SaaS), also known as on-demand software, is a cloud computing model in which a software/application is provided on a subscription basis. The software could literary be any application that supports business operations. With SaaS, resources can be scaled up as the business needs to grow.

Infrastructure as a Service (IaaS) is another set of cloud computing models in which a cloud provider host infrastructure components such as networking hardware, data-center, or servers. Depending on the size and resources of a company, IaaS can be tailored and customized to meet its needs. Once all the required infrastructure has been installed, resources in the cloud can be accessed via a Wide Area Network (WAN) such as the internet. For example, a user can create virtual machines and later install applications used to manage data.

Lastly, Platform as a Service (PaaS) is a computing model in which a cloud provider provides computing components to host applications. Mainly, PaaS provides a framework through which developers that develop and design their applications. For example, a system to validate data can be installed in this model to help validate data running in a SaaS model. It is important to note that these models can be used exclusively or in combination with other models to manage cloud data efficiently.

Recommendation Systems

A recommendation system is an application that can predict user behavior on a given site based on the preferred activities of the user. Recommendation systems came into existence thanks to

technologies such as artificial intelligence, machine learning, and data mining. Recommendation systems are used to recommend items such as books, movies, products, and tourist sites, among many other reasons. It could be a frustrating experience for users to search for a given item and sort them from a list of so many other items. Luckily, recommendation systems can study what users like to do and provide relevant content based on that experience. Recommendation systems require large scale computational and storage resources, which are not always available for young companies. Even for enterprises that have substantial establishments, it is still costly to host a recommendation system due to its flexibility requirements. Therefore recommendation systems have always been offered through cloud infrastructure.

As explained by Fatima (1-6), a recommendation system is a combination of various methods, which primarily include a Data mining system, Machine learning system, and an artificial intelligence system. For users to get the best experience out of these recommendation systems, they must be engaged with contents for a given time. That is, a recommendation system starts its work by mining user data based on their experience browsing through different contents. For example, a user browsing through Action movies will get future action movies recommendation. Once user data has been mined, the machine learning process initiates itself to study user patterns and provide insightful information to an artificial intelligence system. For a recommendation system to recommend content based on the user's behavior, it has to possess some form of intelligence. Meaningful insights from a machine learning system will be fed into an artificial intelligence system to help users reason out and provide suggestions to relevant content.

Application of Cloud-Based Recommendation Systems

Academic Solutions

As the size of content continues to grow on the internet, it becomes apparently tricky for students to find what they need. With a recommendation system, students can get personalized content

hence speeding up research projects and increasing user experience on a site Çano, np). In this domain, recommendation systems are based on the course undertaken by a student and the content searched by the student daily. User-based collaborative filtered, in combination with traditional search systems, are used to realize fully personalized experiences. Most of the recommendation systems operating this area use the Hadoop file system to store a large amount of data. Distributed processing crawlers will then be used to index and cluster user content.

Product Promotion

Marketing is one of the fields where a recommendation system has been widely implemented. This is because a large population of users spent most of their time interacting with different content on the internet (Derick 1-6). Today, a user can visit an e-commerce application to view a specific product of interest. While on another website, a nondisruptive popup will show the exact product that the user had seen. This is an excellent marketing strategy to remind users of the product they had viewed and forgot to purchase.

E-commerce Solution

Recommendation systems are currently being used in e-commerce to increase user experience and also promote other products. A user who searched for a mobile phone with a brand of Nokia will likely get other suggested products made by Nokia. Additionally, a user searching for utensils will probably in the future find other items that are used with those utensils. This is a vital experience for user's who rely on technology and those that expect more from technology.

Current Implementation of Cloud-Based Recommendation System

Google Cloud Prediction API, also known as Cloud Machine learning Engine, is a cloud-based machine learning tool provided to developers in the form of Restful API to them model prediction systems such as spam detection and recommendation system Çano, np). The prediction

system is based on a tensor flow machine learning library. Developers can use this library to extend the functionality of their applications while delivering a fully working cloud-based application.

Amazon Machine Learning is based on the Amazon machine learning platform. This platform makes it easy to create models and generate predictions using an API developed by Amazon. This means that developers won't have to worry about implementing custom prediction code or configuration and management of the required infrastructure. The machine learning technology provided by Amazon is highly scalable and can generate billions of predictions at high throughput.

Azure Machine Learning is a cloud-based machine learning algorithm of Microsoft Azure. It is a set of pre-built packages developed to offer drag and drop gestures. The Azure machine learning platform is available for all data scientist skill levels and provides simple data-flow and time-saving sample experiments.

Oryx is an open-source project offering scalable machine-learning analytic infrastructure. Oryx is a two-tier design consisting of the computational and serving layer. The computational layer is a java-based server layer that builds models using MapReduce and Hadoop. The Server layer exposes a REST API to which developers can connect to. Oryx can be accessed via a web browser or any tool that can make HTTP requests Çano, np).

Conclusion

Because of the flexibility and cost savings it offers, cloud computing is admittedly an imminent computing technology for computational purposes. Today, most applications used by users daily, such as email, file sharing, and multimedia, have been moved in the cloud for accessibility, availability, and convenience purposes. Recommendation systems have been implemented in the cloud to take advantage of big data and flexible resources.

Works Cited

Çano, Erion. "Cloud-based Recommendation Systems: Applications and Solutions." (2020).

Jamiy, Fatima EL, et al. "The potential and challenges of Big data-Recommendation systems next level application." *arXiv preprint arXiv: 1501.03424* (2015) pg1-6.

Leony, Derick, et al. "A cloud-based architecture for an affective recommender system of learning resources." *CEUR-Workshop Proceedings*, 2012 pg1-6.