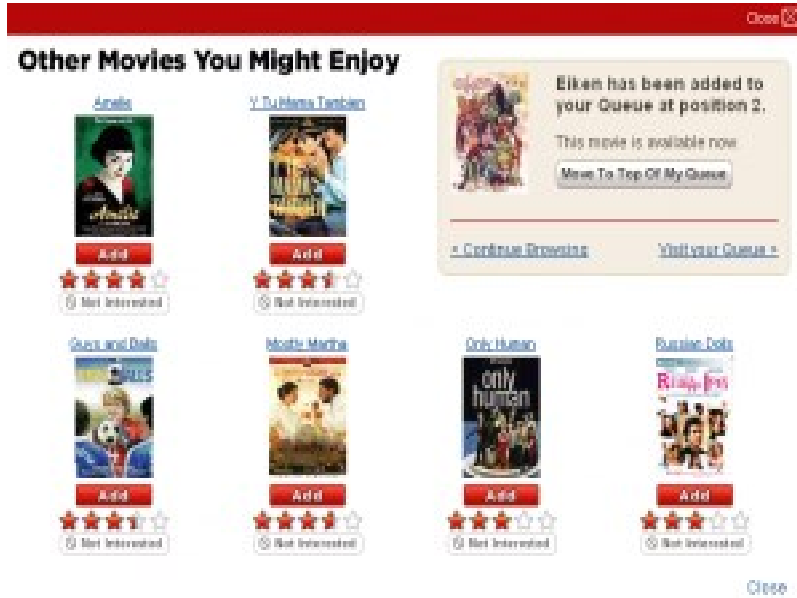


Apache Spark ile Tavsiye Motoru(Recommendation Engine) nasıl yazılır ?

Merhabalar ,

Bu yazımda Tavsiye Motoru(Recommendation Engine) hakkında genel bilgiler vereceğim , sonrasında ise Apache Spark(Java tabanlı) ile geliştirilmiş örnek bir proje paylaşacağım .

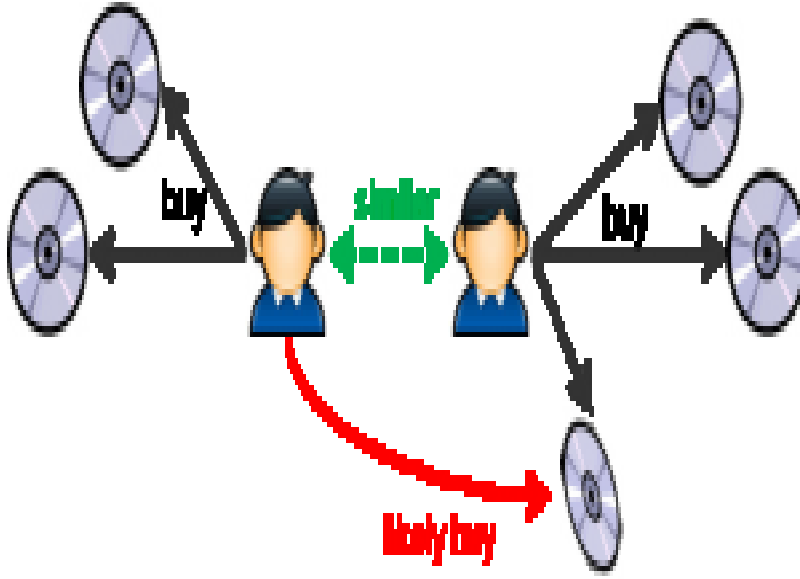
Günlük hayatta web sitelerinde dolaşırken size sürekli öneriler sunulur



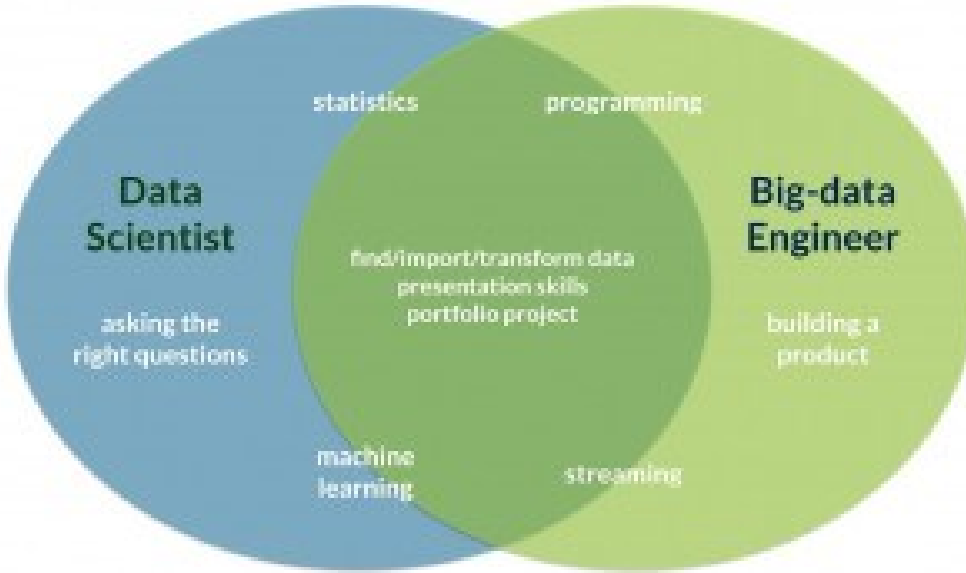
- Sinema içerikli sitelerde ilginizi çekebilecek filmler listesi
- Müzik sitelerinde ilginizi çekebilecek albüm listesi

Genel amaç satış yada gösterim oranlarını artırmaktır.

Aslında tavsiye sistemlerinde genel mantık şu şekildedir . Web siteleri sizin filmlere , kitaplara , müziklere verdiğiniz puanlamaları takip ederek size benzeyen kişiler üzerinden modelleme çkartır . Bu modelleme üzerinden de size bu ürünler tavsiye edilir

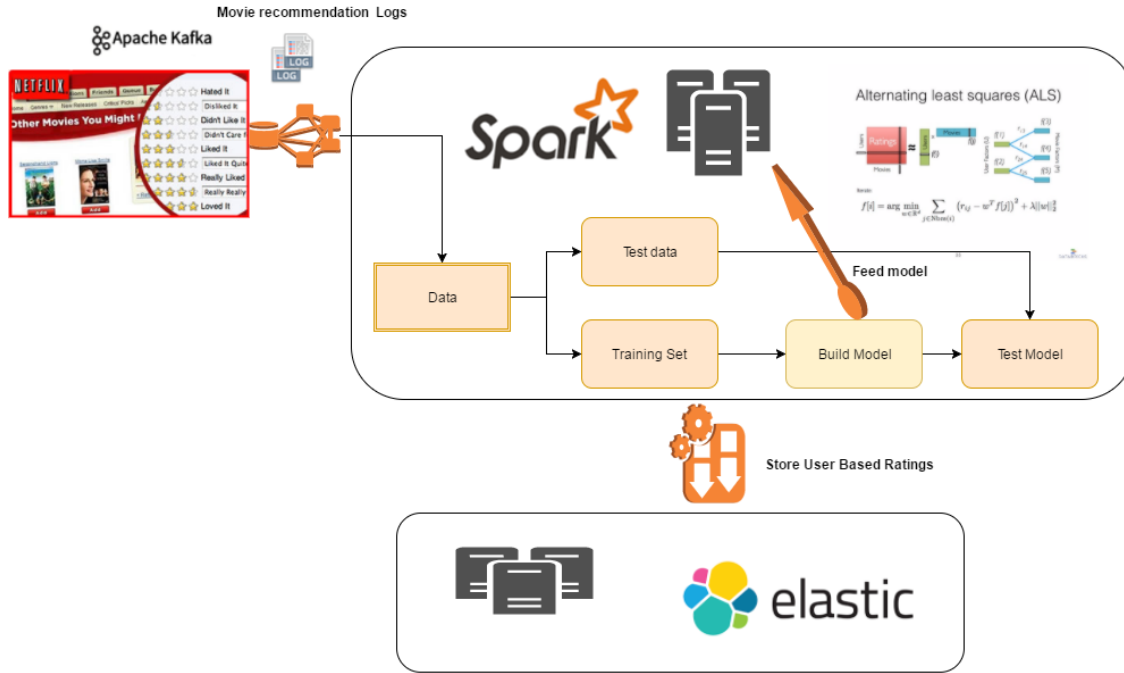


Bu sistemi geliştirebilmek için iki uzmanlık alanında çalışma yapmanız gerekiyor (Big data Engineering + Data Science) . Zaten profesyonel firmalar Big Data ekiplerinde mühendislik geçmişi olan Big Data Engineer ve İstatistik , Matematik geçmişi olan Data Scientist rollerini çalıştırıyorlar .



Tavsiye Motorunu geliştirirken kullandığım sistem ise Apache Spark ve ALS(Alternating Least Squares) algoritması . Aslında ALS algoritması Apache Spark içerisinde build edilmiş bir halde geliyor ve sizin bu algoritmayı tekrar geliştirmenize gerek yok . Sadece nasıl çağrılacağını bilmeniz yeterli .

Genel mimari ;



- Loglar anlık olarak mesajlaşma sistemi üzerinden gelmektedir . Kafka tercih edilebilir fakat yaptığım örnekte logları local dosya sisteminden okudum
- Gelen loglar üzerinden ALS algoritması ile bir model oluşturulur ve test edilebilir
- Oluşturulan model Apache Spark ile tüm kullanıcılar için çalıştırılır ve kullanıcıların beğenebilecekleri filmler çıkartılabilir
- Sonuçlar veritabanı sistemine yazdırılıp , buradan çeşitli sorgulamalar yapılabilir .