Large-Scale Machine Learning: Recommendation Engine for Mobile Apps

Andrew Look
Senior Software Engineer
09/29/2011
Objective

• Incorporate product recommendations into our mobile apps

• Requirements
  – This is an **online system**, meaning that the recommendations evolve as users interact with our webapps
  – Feed cookie/product correlation data into the recommendation engine from our data warehouse
  – Run the recommendation model builder **at scale** (we have lots of data!)
  – Expose a web service endpoint for the mobile apps to talk to
  – Ensure that the web service has a high enough capacity to serve a large number of concurrent users.

• This sounds ambitious, but some power tools are available to help accelerate this process!
Machine Learning? Sounds difficult!

- Machine learning is a challenging area of programming, to be sure
- Libraries exist to implement most of the statistical algorithms for you
  - Mahout is the weapon of choice here
  - Mahout runs on top of Hadoop, so it takes advantage of large-scale distributed cluster processing.
  - Mahout also has a built-in recommendation system, called Taste
    - [https://cwiki.apache.org/confluence/display/MAHOUT/RecommendationExamples](https://cwiki.apache.org/confluence/display/MAHOUT/RecommendationExamples)
    - [https://cwiki.apache.org/MAHOUT/recommender-documentation.html](https://cwiki.apache.org/MAHOUT/recommender-documentation.html)
  - We’ll be using taste to encapsulate the recommendation functionality
- Challenge: gracefully getting data in/out of the Taste Recommender
Taste Recommender Architecture

Web Service

Web log data

Java/J2EE application

Recommender

Neighborhood
  - Correlation
    - Preference Inferrer

DataModel

Preference

User

Item

User

Preference

User

Item

User

Data Store
What do YOU get out of this?

- Work on an application that will be seen by hundreds of thousands of consumers
  - That’s right, if your code works, we’ll push it to production (and give you credit for it!)
- Learning Distributed Computation
  - Hadoop (fascinating, and great to have on your resume)
  - Machine Learning Techniques
- Learning Web Service Performance
  - How do you scale a web service to meet demands of many visitors?
- Learning Mobile Development
  - We’re happy to have you build onto our iOS or Android apps, your choice!
  - How do you manage codebases for different platforms (iPhone/Android)?
  - How do you manage different screen sizes and firmware versions to build resilient applications?
Shopzilla is a Comparison Shopping Engine, with a variety of web applications live throughout the US, UK, France, and Germany

- http://www.bizrate.com
- http://www.bizrate.co.uk
- http://www.beso.com
- http://www.tada.com
- http://www.spardeingeld.de
- ..etc
(Very) Recently, we have built out mobile apps for Bizrate.com on a number of platforms

- iPhone
- iPad
- Android Phone
- Android Tablet

If you’re interested, please check them out!