Why Recommendation Engine

Whenever we have a consumer interacting with some item whether it’s a product, piece of media or something else, there is room for improving the user’s experience by offering other relevant choices in a personalized manner. There is also the potential for increasing the revenue for business in the process.

This is where recommendation engines come into the picture. They provide personalized recommendation based on the user’s behavior and the behavior of other like minded users.

Questions To Ask

As a business owner you need to ask yourself the following questions before choosing a recommendation engine

1. How will make recommendation when I have new customer, with behavioral data available i.e. how will it solve the cold start problem
2. How does it make recommendations when I have very limited user behavior data
3. How does it make recommendations when an user is fully engaged and there is enough behavioral data.
4. What if my users do not explicitly rate items
5. Can it make recommendations in real time, based on user’s current behavior
6. Some times recommendation are too expected and there is no element of surprise. How does the system handle such scenario
7. What kind of input data is needed
8. How configurable is the system
9. What if I have some business goal for my products and how do they reconcile with recommendations
10. How does the system scale with increasing load
11. How seamlessly does it integrate with exiting IT infrastructure
12. Is it cloud ready
13. Is there a convenient licensing model
14. Once deployed, how do I know that the system is working and I have ROI

Sifarish Has the Answers

Sifarish is an open source recommendation engine running on a Big Data platform comprising of *Apache* *Hadoop*, *Apache* *Storm* and *Redis*. *Sifarish* has satisfactory answer to all the issues raised above as outlined below.

Cold start problem

For brand new customer, *sifarish* does matching between user profile and product attributes. If user profile is not available, *sifarish* will use popular items to make recommendation. In other words, sifarish will bootstrap new customers one way or other. These solutions are based on geometric algorithms. These solutions are implemented on Hadoop

Warm start problem

When an user has engaged with very few item, sifarish will match those items with similar items to make recommendations. Again geometric algorithms are used for the matching purpose. These solutions are implemented on Hadoop.

These computations need to be repeated periodically as an user gets exposed to engages with new items.

Fully engaged users

When users are fully engaged and enough behavioral data is available, sifarish uses state of the art collaborative filtering algorithms to make recommendations. These solutions are implemented on Hadoop.

The solution is based on all users past behavioral data over a time window spanning weeks or months as per how it’s configured.

User rating

Even if explicit user ratings for items are not available, sifarish can process click stream data to evaluate user’s affinity for different items and estimate implicit rating. User ratings data is the key input for collaborative filtering algorithms.

Real time recommendation

*Sifarish* offers real time recommendations based on users current behavior in a predefined small time window. The solution is based on Apache Storm. It also uses Hadoop to compute item correlation matrix based on historical user behavior data.

Diversity and Novelty

*Sifarish* uses sophisticated algorithms to increase diversity, novelty and controversiality in the recommendation mix. The amount of diversity and novelty to be introduced can be set at per user level.

Input Data

For cold start and problems, *sifarish* needs item or product data. In addition, user profile data could also be used if available. *Sifarish* supports various attribute types including numerical, categorical, text, semantic,, time and location..

For fully engaged users, *sifarish* uses click stream data as input to it’s collaborative filtering algorithms.

Extreme configurability

*Sifarish* believes in giving power in the hands of the end users. There are many configuration parameters to control the behavior of the solution. This makes it possible to male sifarish effective in any situation depending on the nature of business and nature of the customer population.

Business goal

Some times there may be business goals associated with items (e.g., promoting items with high inventory etc) which may be at odds with recommendations. *Sifarish* provides a way to reconcile recommendations with business interest by using relative weighting.

Scalabilty

Sifarish runs on an open source Big Data platform consisting of Hadoop, Storm and Redis. This platform is horizontally scalable with commodity machines.

Easy integration

Sifarish stays away from complex file format and uses CSV file format for input and output. As long as the external enterprise system can generate CSV output to be consumed by sifarish and can consume CSV data generated by sifarish, there is no additional work necessary for integration.

Additional more tighter integration adapters can be easily be built. For example recommendation output from *Hadoop* can imported to enterprise system RDBMS directly.

Ready for cloud

Sifarish can easily be deployed on Amazon or Google cloud. This approach makes more sense if the existing enterprise solution is already hosted on cloud.

Licensing model

Sifarish has dual model. The community edition is free with software, scripts, tutorial and plenty of blogs for background technical material. The enterprise edition includes web based admit tool, support, training , consultancy and documentation.

Measuring effectiveness

*Sifarish* comes with some Hadoop based analytical tools to measure and track effectiveness of recommendations