

# Factorization Machine (FM)

Content Prepared By: Chandra Lingam, Cotton Cola Designs LLC

Copyright © 2017 Cotton Cola Designs LLC. All Rights Reserved.

All other registered trademarks and/or copyright material are of their respective owners

# Factorization Machine

Factorization Machine algorithm is optimized for handling high dimensional sparse datasets

Supports Regression and Classification

Personalize Content - “predict” ratings/likeness

- Click Prediction for Ad-Placement

- Product recommendation for user

- Movie recommendation

- News/Social Media Feed personalization for users

# Factorization Machines

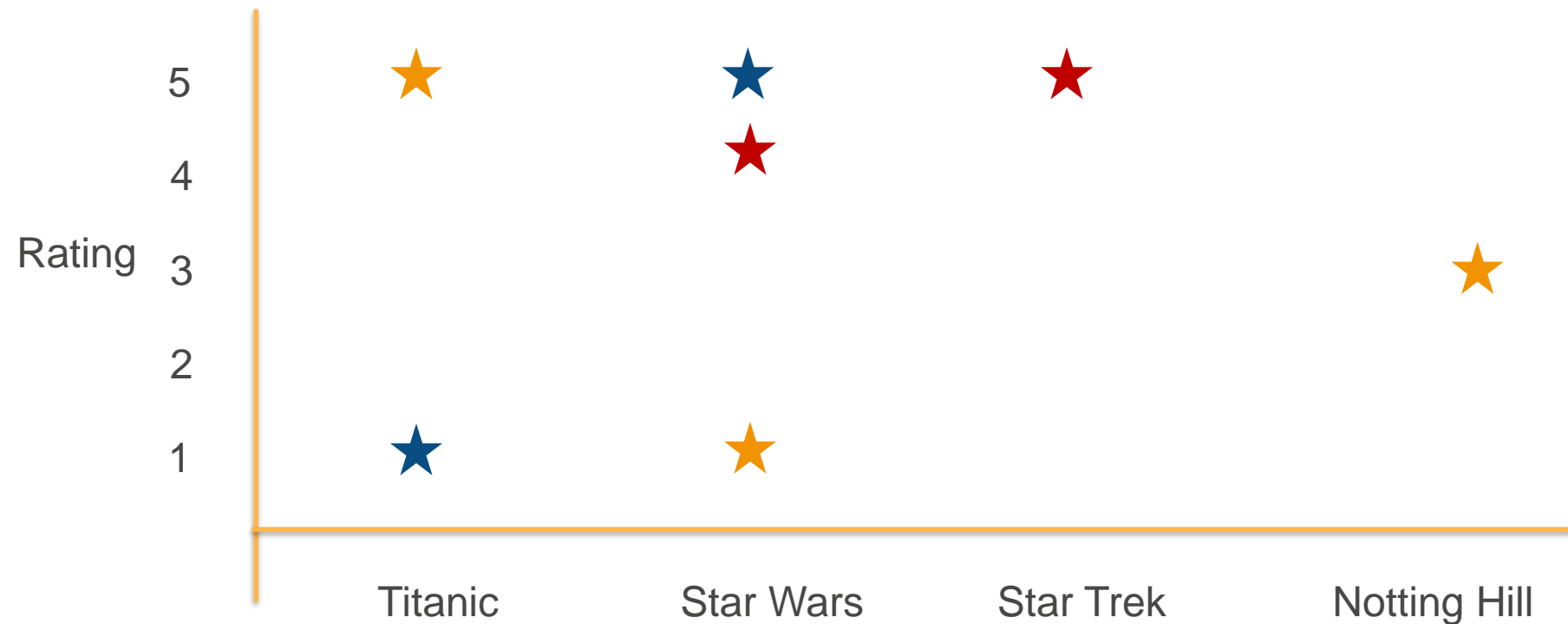
Models all interactions between features using Factorized Parameters

Estimate interactions with very sparse datasets

Linear complexity for computing model parameters

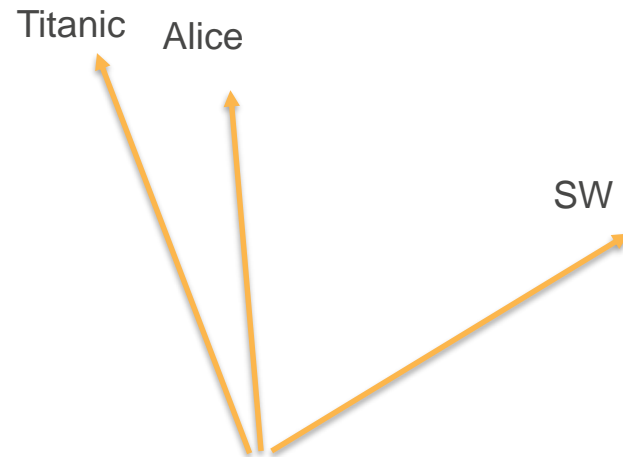
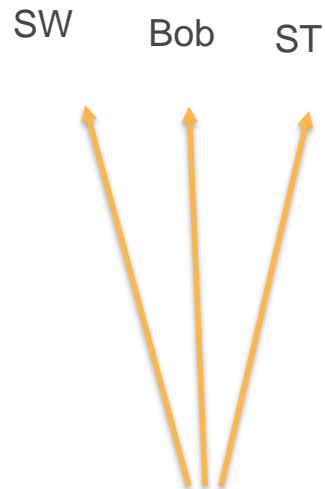
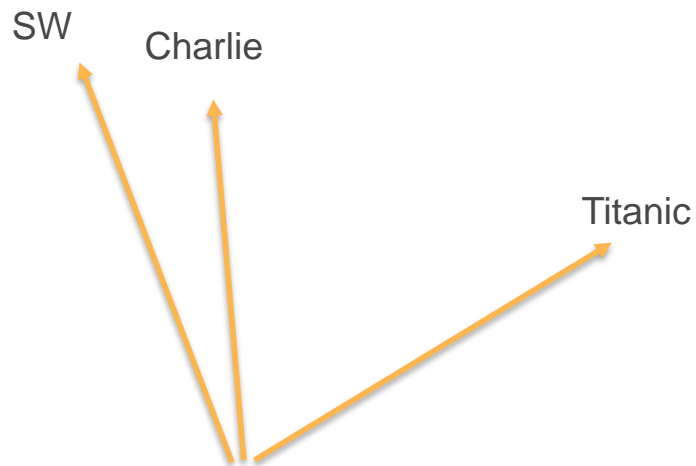
Supports very large datasets

# Movie and User

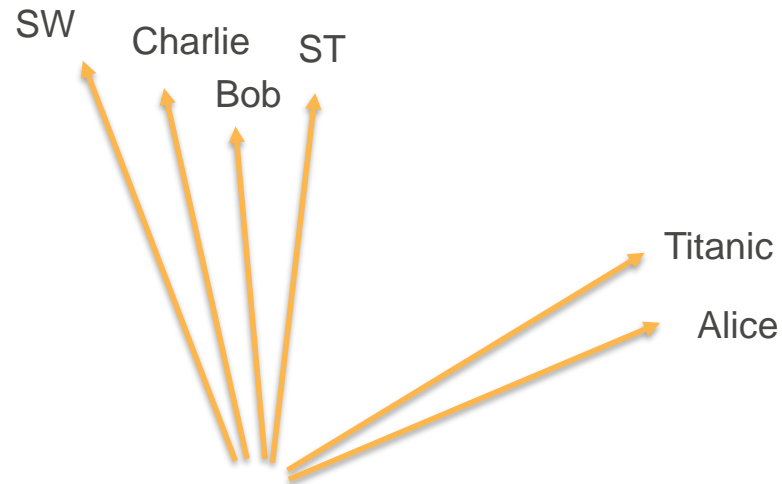


★ Alice    ★ Bob    ★ Charlie

# Pair Wise Interaction



# Recommendation



# Factorization Machine – Data Format

Input:

recordio-protobuf (with Float32 values)

Inference:

json

recordio-protobuf

# Demo – Movie Recommendation

Movie Lens [Dataset](#)

Predict how a user would rate a movie

Recommend movies based on user rating, other similar users and other similar movies

fm\movie\_data\_preparation.ipynb,  
fm\_cloud\_training\_template.ipynb,  
fm\_cloud\_prediction\_template.ipynb



# Demo Movie Recommendation Files

File Name	Purpose
Movies.csv	List of movies [movie id, title, genre]
Ratings.csv	Movies ratings by user [user id, movie id, rating]
Movie_genre.csv	Movies with Genre in separate columns
user_movie_{train test}.recordio	Sparse RecordIO Train/Test Data – OneHotEncoded [user id, movie id], Rating
user_movie_{train test}.svm	Sparse SVM Train/Test Data – Easy to read with text editor.
one_hot_enc_movies.svm	List of movie ids and corresponding one hot encoded movie column identifier
one_hot_enc_users.svm	List of user ids and corresponding one hot encoded user column identifier

# Useful Resources

[Factorization Machines](#) by Steffen Rendle

[LibFM](#) Software

[Comparison of LibFM Implementations](#) by Alex Rogozhnikov

[Collaborative Filtering](#) by Anand Rajaraman