

# Word Docs Summary

---

*A very simple point by point summary of the topics of the exam of Recommender System. It may be not complete, because professor Cremonesi has talked about a few other topics (not many!) during his lectures, this is just the list of the chapters from his material on BeeP.*

*Edited by: William Bonvini*

## Word Docs Summary

### Basic 1

- Introduction**
- Taxonomy**
- Ratings, predictions and recommendations**
- Inferring preferences**
- Non-personalized recommenders**
- Ratings, predictions and recommendations**
- Global effects**

### Basic 2

- Requirements**
- Quality Indicators**
- Evaluation Techniques**
- Offline Evaluation**
- Algorithms for RSs**
- Evaluation Techniques in depth**
- Error Metrics**
- Classification Metrics**
- Combining Metrics**
- Ranking Metrics**
- Evaluating Diversity**
- Evaluating Novelty**

### Basic 3

- Content Based Filtering**
- Cosine Similarity**
- Estimating Ratings**
- Similarity Matrix**
- Improving the Item-Content Matrix**
- TF-IDF**
- Users Based Filtering**

### Basic 4

- Collaborative User Based**
- Item Based Collaborative Filtering**
- Some Considerations about CF**
- Recommendation as Association Rules**

### Advanced 1

- Estimating ratings as an optimization problem**
- 2-norm vs F-norm**
- Missing as Negative vs Missing as Random**
- Overfitting and the regularization of the model**
- Stochastic Gradient Descent**
- Learning to rank**
- BPR**

### Advanced 2

- Matrix Factorization**
- SVD ++**

## **Model Based Matrix Factorization**

Advanced 3

### **Hybrid Recommender Systems**

#### **S SLIM**

### **Context Aware Recommender Systems**

Advanced 4

### **Factorization Machines Introduction**

#### **Explaining FM's formula**

#### **Comparison to CF techniques**

#### **Benefits and harms of FMs**

#### **Mixing different kinds of filtering**

# Basic 1

---

## Introduction

- Recommend Systems
- Data Representation

## Taxonomy

- categorizing algorithms
- personalized recommenders
- collaborative filtering

## Ratings, predictions and recommendations

### Inferring preferences

- formal representation of INPUT
- sparsity of URM

### Non-personalized recommenders

## Ratings, predictions and recommendations

- rating distribution
- computing ratings

### Global effects

- general description
- computing averages and biases

# Basic 2

---

## Requirements

- Functional
- Non Functional
  - Response Time
  - Scalability
  - Privacy and Security
  - User Interface

## Quality Indicators

- Consistency
- Confidence
- Coverage
- Diversity
- Novelty
- Serendipity
- Relevance

## Evaluation Techniques

- Direct Users' Feedbacks
- A/B Testing
- Controlled Experiment
- Crowdsourcing

## Offline Evaluation

- Ground Truth
- Top-K Recommendations
- Error metrics
- Classification metrics
- Algorithm

## Algorithms for RSs

- The recommender algorithm
- Relevant Data Sets
- Model Based Algorithms, Memory Based Algorithms

## Evaluation Techniques in depth

- Hold-out technique
- the Netflix Prize
- K-fold evaluation
- LOO technique

## Error Metrics

- MAE
- RMSE
- limitations of error metrics
- comparing the distributions of metrics

## **Classification Metrics**

- Recall
- Precision
- the implicit assumption
- missing ratings

## **Combining Metrics**

- Precision or Recall?
- F-Measure
- ROC curve
- Popularity Bias
- Utility

## **Ranking Metrics**

- Ranking in general
- ARHR
- MAP
- MAP's interpretation
- Spearman's Rho
- Kendall's Tau

## **Evaluating Diversity**

- defining diversity
- measuring the distance
- balancing quality and diversity

## **Evaluating Novelty**

- defining novelty
- estimating novelty

# Basic 3

---

## Content Based Filtering

- CBF Definition
- ICM

## Cosine Similarity

- Measuring Similarity
- The Cosine
- Shrinking

## Estimating Ratings

- Estimating

## Similarity Matrix

- Definition
- the choice of  $K$

## Improving the Item-Content Matrix

- Non Binary attributes
- Attribute weights

## TF-IDF

- Definition
- Term Frequency
- Inverse Document Frequency

## Users Based Filtering

- Definition

# Basic 4

---

## Collaborative User Based

- URM
- User Similarity
- Pearson Coefficient
- the delta

## Item Based Collaborative Filtering

- IBCF
- how to compute similarity on explicit ratings
- adjusted cosine

## Some Considerations about CF

- Normalization
- KNN
- choosing between user based and item based CF

## Recommendation as Association Rules

- the point of view of association rules

# Advanced 1

---

## Estimating ratings as an optimization problem

- recap in IBCF
- minimizing the estimation error
- the optimization problem setting

## 2-norm vs F-norm

- Frobenius Norm

## Missing as Negative vs Missing as Random

- the choice of the assumption
- Frobenius norm's problem
- Quality Metrics for ML approaches

## Overfitting and the regularization of the model

- the risk of overfitting
- the regularization term
- the problems of machine learning approaches

## Stochastic Gradient Descent

- definition
- the gradient descent problem and the learning rate
- implementing a Stochastic Gradient Descent

## Learning to rank

- the learning to Rank approach
- ranking error metrics
  - list-wise metric
  - point-wise metric
  - pair-wise metric

## BPR

- definition
- stochastic gradient descent
- the risk with the gradient
- BPR with implicit ratings
- BPR error function
  - regularization term!
- BPR with explicit ratings



# Advanced 2

---

## Matrix Factorization

- introduction
- FunkSVD
- Latent Factors
- Avoiding Overfitting

## SVD ++

- definition
- the assumption on zero elements

## Model Based Matrix Factorization

- Adapting SVD++
- matrix formulation
- Asymmetric SVD

# Advanced 3

---

## Hybrid Recommender Systems

- Definition
- Linear Combination
- disadvantages of linear combination
- Combining Lists
- Pipelining Algorithms
- pipeline disadvantages
- Merging Models

## S SLIM

- Definition
- Weighted Average
- the idea of S SLIM
- how to choose the value of alpha

## Context Aware Recommender Systems

- definition
- tensor factorization
- how to learn the three matrices
- limitations of tensor factorization

# Advanced 4

---

## Factorization Machines Introduction

- FM Definition
- Matrix Representation
- FM formulation

## Explaining FM's formula

- Constant component
- Vector component
- improved global effects
- matrix component
- handling big numbers of parameters

## Comparison to CF techniques

- Similarity to CF with matrix factorization
- Why FMs are better than classical CF techniques

## Benefits and harms of FMs

- FMs flexibility, an example
- FMs issues

## Mixing different kinds of filtering

- putting together collaborative and content based
- the general problem of content-based predictions
- balancing content-based parameters

## Solving the optimization problem

- implementing the algorithm
- the classification problem
- quality of the input