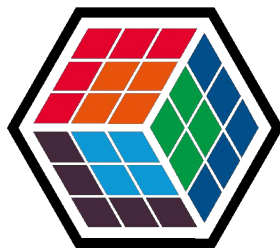


THE DEVELOPER'S CONFERENCE

Trilha – Java

Kamila Santos Oliveira
Software Developer



THE DEVELOPER'S CONFERENCE

IA com Java?

Sistema de recomendação em Java com o Apache Mahout

Agenda



THE
DEVELOPER'S
CONFERENCE

- ❖ O que é machine learning?
- ❖ O que são sistemas de recomendação?
- ❖ Person Correlation Similarity
- ❖ Como funciona?
- ❖ VAMOS AO CÓDIGO?



THE
DEVELOPER'S
CONFERENCE

O que é machine learning?

Se trata do aprendizado de máquina, um ramo da IA (inteligência Artificial) que se baseia na ideia de que sistemas são capazes de aprender com dados, identificar padrões e realizar tomadas de decisões com o mínimo de intervenção de um ser humano.

Machine Learning

O que são sistemas de recomendação?

São técnicas de software que fornecem sugestões de itens para ajudar usuários deste sistema nos diversos processos de tomada de decisão.



THE
DEVELOPER'S
CONFERENCE

Sistemas de recomendação



THE
DEVELOPER'S
CONFERENCE

Person Correlation Similarity

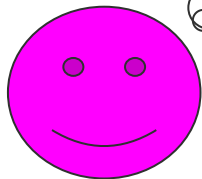
Se baseia na similaridade entre pessoas na hora de escolher e atribuir notas para determinar os resultados de sua recomendação.

**Person
Correlation
Similarity**



THE
DEVELOPER'S
CONFERENCE

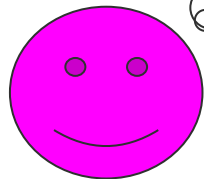
QUE FILME EU
VOU ASSISTIR?



Como
funciona?



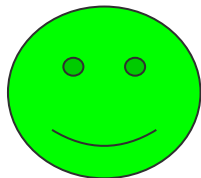
QUE FILME EU
VOU ASSISTIR?



FILME A	NOTA 8
FILME B	NOTA 1
FILME C	NOTA 7



PESSOA 1 : JÁ ASSISTI
ESTES AQUI



FILME A	NOTA 8
FILME D	NOTA 10
FILME E	NOTA 5



PESSOA 2 : JÁ ASSISTI
ESTES AQUI



FILME A	NOTA 2
FILME B	NOTA 1
FILME E	NOTA 3



THE
DEVELOPER'S
CONFERENCE

Configuração POM.xml



THE
DEVELOPER'S
CONFERENCE

Onde ficam as dependências
utilizadas no nosso projeto.

Vamos adicionar o Apache Mahout e
o Hadoop Client



THE
DEVELOPER'S
CONFERENCE

Apache Mahout



MAHOUT



For Creating Scalable Performant
Machine Learning Applications

 [Download Mahout](#)

Currently v0.14.0



THE
DEVELOPER'S
CONFERENCE

Apache Hadoop



Apache Hadoop

The Apache™ Hadoop® project develops open-source software for reliable, scalable, distributed computing.

The Apache Hadoop software library is a framework that allows for the distributed processing of large data sets across clusters of computers using simple programming models. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage. Rather than rely on hardware to deliver high-availability, the library itself is designed to detect and handle failures at the application layer, so delivering a highly-available service on top of a cluster of computers, each of which may be prone to failures.

[Learn more »](#)

[Download »](#)

[Getting started »](#)

Latest news

[Apache Hadoop](#)

2019 May 7

Modules

The project includes these modules:

Related projects

Other Hadoop-related projects at Apache include:



THE DEVELOPER'S CONFERENCE

```
1 <project xmlns="http://maven.apache.org/POM/4.0.0" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
2   xsi:schemaLocation="http://maven.apache.org/POM/4.0.0 http://maven.apache.org/maven-v4_0_0.xsd">
3   <modelVersion>4.0.0</modelVersion>
4   <groupId>br.com.kamila</groupId>
5   <artifactId>systema-de-recomendacao-java-mahout</artifactId>
6   <packaging>jar</packaging>
7   <version>1.0-SNAPSHOT</version>
8   <name>systema-de-recomendacao-java-mahout</name>
9   <url>http://maven.apache.org</url>
10  <dependencies>
11    <dependency>
12      <groupId>junit</groupId>
13      <artifactId>junit</artifactId>
14      <version>3.8.1</version>
15      <scope>test</scope>
16    </dependency>
17    <dependency>
18      <groupId>org.apache.mahout</groupId>
19      <artifactId>mahout-mr</artifactId>
20      <version>0.12.2</version>
21    </dependency>
```



```
20     </version>0.12.2</version>
21   </dependency>
22   <dependency>
23     <groupId>org.apache.hadoop</groupId>
24     <artifactId>hadoop-client</artifactId>
25     <version>2.7.2</version>
26   </dependency>
27 </dependencies>
28 </project>
29
```



THE
DEVELOPER'S
CONFERENCE

Conteúdo do CSV



THE DEVELOPER'S CONFERENCE

src ▾ main ▾ resources ▾ filmes.csv

```
61195 615,231,9
61196 1264,135,10
61197 2719,28,10
61198 4109,49,10
61199 2587,344,9
61200 5859,174,10
61201 3363,13,8
61202 686,284,10
61203 1690,344,10
61204 542,221,10
```



THE
DEVELOPER'S
CONFERENCE

Recomendador.java



```
src ▸ main ▸ java ▸ br ▸ com ▸ kamila ▸ Recomendador.java ▸ Recomendador ▸ getModeloDeFilmes()
```

```
1 package br.com.kamila;
2
3 import org.apache.mahout.cf.taste.impl.model.file.FileDataModel;
4 import org.apache.mahout.cf.taste.model.DataModel;
5
6 import java.io.File;
7 import java.io.IOException;
8 public class Recomendador {
9
10     public DataModel getModeloLivros() throws IOException {
11         return getModelo("livros.csv");
12     }
13
14     private DataModel getModelo(String path) throws IOException {
15         File file = new File("src/main/resources/" + path);
16         return new FileDataModel(file);
17     }
18
19     public DataModel getModeloDeFilmes() throws IOException {
20         return getModelo("filmes.csv");
21     }
22 }
```



THE
DEVELOPER'S
CONFERENCE

RecomendadorBuilder. java



src ▸ main ▸ java ▸ br ▸ com ▸ kamila ▸ RecomendadorBuilder.java ▸ RecomendadorBuilder ▸ buildRecommender(DataModel)

```
1 package br.com.kamila;
2
3 import org.apache.mahout.cf.taste.common.TasteException;
4 import org.apache.mahout.cf.taste.eval.RecommenderBuilder;
5 import org.apache.mahout.cf.taste.impl.neighborhood.ThresholdUserNeighborhood;
6 import org.apache.mahout.cf.taste.impl.recommender.GenericUserBasedRecommender;
7 import org.apache.mahout.cf.taste.impl.similarity.PearsonCorrelationSimilarity;
8 import org.apache.mahout.cf.taste.model.DataModel;
9 import org.apache.mahout.cf.taste.neighborhood.UserNeighborhood;
10 import org.apache.mahout.cf.taste.recommender.Recommender;
11 import org.apache.mahout.cf.taste.recommender.UserBasedRecommender;
12 import org.apache.mahout.cf.taste.similarity.UserSimilarity;
13 public class RecomendadorBuilder implements RecommenderBuilder {
14     public Recommender buildRecommender(DataModel model) throws TasteException {
15         UserSimilarity similarity = new PearsonCorrelationSimilarity(model);
16         UserNeighborhood neighborhood = new ThresholdUserNeighborhood(0.1, similarity, model);
17         UserBasedRecommender recommender = new GenericUserBasedRecommender(model, neighborhood, similarity);
18         return recommender;
19     }
20 }
```




THE
DEVELOPER'S
CONFERENCE

RecomendaFilmes.java



```
src ▾ main ▾ java ▾ br ▾ com ▾ kamila ▾ Recomendafilmes.java ▾ Recomendafilmes ▾ main(String[])
6   import org.apache.mahout.cf.taste.recommender.Recommender;
7
8   import java.io.IOException;
9   import java.util.List;
10
11  public class Recomendafilmes {
12      Run | Debug
13      public static void main(String[] args) throws IOException, TasteException {
14          DataModel filmes = new Recomendador().getModeloDeFilmes();
15          Recommender recommender = new RecomendadorBuilder().buildRecommender(filmes);
16          List<RecommendedItem> recommendations = recommender.recommend(8, 3);
17          for (RecommendedItem recommendation : recommendations) {
18              System.out.println("Voce pode gostar deste filme");
19              System.out.println(recommendation);
20          }
21      }
22  }
23 }
```



THE DEVELOPER'S CONFERENCE

```
[main] 17:08:01 INFO org.apache.mahout.cf.taste.impl.model.file.FileDataModel.<init>(FileDataModel.java:185)-Creating
FileDataModel for file src\main\resources\filmes.csv
[main] 17:08:01 INFO org.apache.mahout.cf.taste.impl.model.file.FileDataModel.processFile(FileDataModel.java:357)-Read
ing file info...
[main] 17:08:02 INFO org.apache.mahout.cf.taste.impl.model.file.FileDataModel.processFile(FileDataModel.java:368)-Read
lines: 80710
[main] 17:08:02 INFO org.apache.mahout.cf.taste.impl.model.GenericDataModel.<init>(GenericDataModel.java:113)-Processe
d 7118 users
[main] 17:08:02 DEBUG org.apache.mahout.cf.taste.impl.recommender.GenericUserBasedRecommender.recommend(GenericUserBase
dRecommender.java:86)-Recommending items for user ID '8'
[main] 17:08:02 DEBUG org.apache.mahout.cf.taste.impl.recommender.GenericUserBasedRecommender.recommend(GenericUserBase
dRecommender.java:101)-Recommendations are: [RecommendedItem[item:429, value:10.0], RecommendedItem[item:428, value:10.
0], RecommendedItem[item:392, value:9.932595], RecommendedItem[item:441, value:9.737432], RecommendedItem[item:438, val
ue:9.645442]]
Voce pode gostar deste filme
RecommendedItem[item:429, value:10.0]
Voce pode gostar deste filme
RecommendedItem[item:428, value:10.0]
Voce pode gostar deste filme
RecommendedItem[item:392, value:9.932595]
Voce pode gostar deste filme
RecommendedItem[item:441, value:9.737432]
Voce pode gostar deste filme
RecommendedItem[item:438, value:9.645442]
```



THE
DEVELOPER'S
CONFERENCE

RecomendaLivros.java



THE DEVELOPER'S CONFERENCE

```
src ▸ main ▸ java ▸ br ▸ com ▸ kamila ▸ RecomendaLivros.java ▸ RecomendaLivros ▸ main(String[])
1  package br.com.kamila;
2
3  import org.apache.mahout.cf.taste.common.TasteException;
4  import org.apache.mahout.cf.taste.model.DataModel;
5  import org.apache.mahout.cf.taste.recommender.RecommendedItem;
6  import org.apache.mahout.cf.taste.recommender.Recommender;
7
8  import java.io.IOException;
9  import java.util.List;
10
11 public class RecomendaLivros {
12
13     Run | Debug
14     public static void main(String[] args) throws TasteException, IOException {
15         DataModel produtos = new Recomendador().getModeloLivros();
16         Recommender recommender = new RecomendadorBuilder().buildRecommender(produtos);
17         System.out.println("usuario 1");
18         List<RecommendedItem> recommendations = recommender.recommend(1, 4);
19         for (RecommendedItem recommendation : recommendations) {
20             System.out.println("usuário 1, estes livros combinam com voce");
21             System.out.println(recommendation);
22         }
23
24         System.out.println("usuario 2");
25         recommendations = recommender.recommend(2, 4);
26         for (RecommendedItem recommendation : recommendations) {
27             System.out.println("usuário 2, estes livros combinam com voce");
28             System.out.println(recommendation);
29         }
30     }
31 }
```



THE DEVELOPER'S CONFERENCE

```
[main] 17:21:13 INFO org.apache.mahout.cf.taste.impl.model.file.FileDataModel.<init>(FileDataModel.java:185)-Creating
FileDataModel for file src\main\resources\livros.csv
[main] 17:21:13 INFO org.apache.mahout.cf.taste.impl.model.file.FileDataModel.processFile(FileDataModel.java:357)-Reading
file info...
[main] 17:21:13 INFO org.apache.mahout.cf.taste.impl.model.file.FileDataModel.processFile(FileDataModel.java:368)-Reading
lines: 126676
[main] 17:21:13 INFO org.apache.mahout.cf.taste.impl.model.GenericDataModel.<init>(GenericDataModel.java:113)-Processed
7118 users
usuario 1
[main] 17:21:13 DEBUG org.apache.mahout.cf.taste.impl.recommender.GenericUserBasedRecommender.recommend(GenericUserBased
Recommender.java:86)-Recommending items for user ID '1'
[main] 17:21:13 DEBUG org.apache.mahout.cf.taste.impl.recommender.GenericUserBasedRecommender.recommend(GenericUserBased
Recommender.java:101)-Recommendations are: [RecommendedItem[item:426, value:9.698958], RecommendedItem[item:428, value:9.595195],
RecommendedItem[item:429, value:9.595195], RecommendedItem[item:439, value:9.544896]]
usuário 1, estes livros combinam com voce
RecommendedItem[item:426, value:9.698958]
usuário 1, estes livros combinam com voce
RecommendedItem[item:428, value:9.595195]
usuário 1, estes livros combinam com voce
RecommendedItem[item:429, value:9.595195]
usuário 1, estes livros combinam com voce
RecommendedItem[item:439, value:9.544896]
usuário 2
```



```
usuario 2
```

```
[main] 17:21:13 DEBUG org.apache.mahout.cf.taste.impl.recommender.GenericUserBasedRecommender.recommend(GenericUserBasedRecommender.java:86)-Recommending items for user ID '2'
```

```
[main] 17:21:13 DEBUG org.apache.mahout.cf.taste.impl.recommender.GenericUserBasedRecommender.recommend(GenericUserBasedRecommender.java:101)-Recommendations are: [RecommendedItem[item:339, value:10.0], RecommendedItem[item:438, value:9.727503], RecommendedItem[item:439, value:9.601675], RecommendedItem[item:412, value:9.581487]]
```

```
usuário 2, estes livros combinam com voce
```

```
RecommendedItem[item:339, value:10.0]
```

```
usuário 2, estes livros combinam com voce
```

```
RecommendedItem[item:438, value:9.727503]
```

```
usuário 2, estes livros combinam com voce
```

```
RecommendedItem[item:439, value:9.601675]
```

```
usuário 2, estes livros combinam com voce
```

```
RecommendedItem[item:412, value:9.581487]
```



THE
DEVELOPER'S
CONFERENCE

Avaliador.java



src ▸ main ▸ java ▸ br ▸ com ▸ kamila ▸ Avaliador.java ▸ Avaliador ▸ main(String[])

```
10 import java.io.IOException;
11 public class Avaliador {
    Run | Debug
12     public static void main(String[] args) throws IOException, TasteException {
13         RandomUtils.useTestSeed();
14         DataModel modelo = new Recomendador().getModeloDeFilmes();
15         RecommenderEvaluator evaluator = new AverageAbsoluteDifferenceRecommenderEvaluator();
16         RecommenderBuilder builder = new RecomendadorBuilder();
17         double erro = evaluator.evaluate(builder, null, modelo, 0.9, 1.0);
18         System.out.println("A taxa de erro desse sistema de recomendacao atualmente e:");
19         System.out.println(erro);
20     }
21 }
22 }
23 }
```



```
[main] 17:27:45 INFO org.apache.mahout.cf.taste.impl.model.file.FileDataModel.<init>(FileDataModel.java:185)-Creating
FileDataModel for file src\main\resources\filmes.csv
[main] 17:27:45 INFO org.apache.mahout.cf.taste.impl.model.file.FileDataModel.processFile(FileDataModel.java:357)-Read
ing file info...
[main] 17:27:46 INFO org.apache.mahout.cf.taste.impl.model.file.FileDataModel.processFile(FileDataModel.java:368)-Read
lines: 80710
[main] 17:27:46 INFO org.apache.mahout.cf.taste.impl.model.GenericDataModel.<init>(GenericDataModel.java:113)-Processe
d 7118 users
[main] 17:27:46 INFO org.apache.mahout.cf.taste.impl.eval.AbstractDifferenceRecommenderEvaluator.evaluate(AbstractDiff
erenceRecommenderEvaluator.java:105)-Beginning evaluation using 0.9 of FileDataModel[dataFile:c:\Users\724354\Documents
\recomendador\istema-de-recomendacao-java-mahout\src\main\resources\filmes.csv]
[main] 17:27:46 INFO org.apache.mahout.cf.taste.impl.model.GenericDataModel.<init>(GenericDataModel.java:113)-Processe
d 6983 users
[main] 17:27:46 INFO org.apache.mahout.cf.taste.impl.eval.AbstractDifferenceRecommenderEvaluator.getEvaluation(Abstrac
tDifferenceRecommenderEvaluator.java:181)-Beginning evaluation of 3443 users
```



THE DEVELOPER'S CONFERENCE

```
tDifferenceRecommenderEvaluator.java:181)-Beginning evaluation of 3443 users
[main] 17:27:46 INFO org.apache.mahout.cf.taste.impl.eval.AbstractDifferenceRecommenderEvaluator.execute(AbstractDifferenceRecommenderEvaluator.java:194)-Starting timing of 3443 tasks in 8 threads
[pool-1-thread-1] 17:27:46 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:55)-Average time per recommendation: 24ms
[pool-1-thread-1] 17:27:46 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:58)-Approximate memory used: 15MB / 126MB
[pool-1-thread-1] 17:27:46 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:59)-Unable to recommend in 0 cases
[pool-1-thread-7] 17:27:47 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:55)-Average time per recommendation: 8ms
[pool-1-thread-7] 17:27:47 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:58)-Approximate memory used: 266MB / 364MB
[pool-1-thread-7] 17:27:47 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:59)-Unable to recommend in 74 cases
[pool-1-thread-2] 17:27:47 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:55)-Average time per recommendation: 6ms
[pool-1-thread-2] 17:27:47 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:58)-Approximate memory used: 210MB / 525MB
[pool-1-thread-2] 17:27:47 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:59)-Unable to recommend in 122 cases
[pool-1-thread-3] 17:27:48 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:55)-Average time per recommendation: 5ms
[pool-1-thread-3] 17:27:48 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:58)-Approximate memory used: 115MB / 525MB
[pool-1-thread-3] 17:27:48 INFO org.apache.mahout.cf.taste.impl.eval.StatsCallable.call(StatsCallable.java:59)-Unable to recommend in 264 cases
```



THE
DEVELOPER'S
CONFERENCE

```
[main] 17:27:48 INFO org.apache.mahout.cf.taste.impl.eval.AbstractDifferenceRecommenderEvaluator.evaluate(AbstractDifferenceRecommenderEvaluator.java:127)-Evaluation result: 1.1215770863180203
```

A taxa de erro desse sistema de recomendacao atualmente e:

```
1.1215770863180203
```

>

Referências:



THE
DEVELOPER'S
CONFERENCE

<https://github.com/Kamilahsantos/sistema-de-recomendacao-java-mahout>

<https://cursos.alura.com.br/course/machine-learning-introducao-aos-sistemas-de-recomendacoes>

https://www.sas.com/pt_br/insights/analytics/machine-learning.html

<https://king.host/blog/2018/09/o-que-sao-sistemas-de-recomendacao/>

<https://www.infoq.com/br/presentations/machine-learning-em-java-com-apache-mahout>

https://www.tutorialspoint.com/mahout/mahout_machine_learning.htm

<https://www.devmedia.com.br/introducao-ao-maven/25128>

<https://mahout.apache.org/>

<https://mvnrepository.com/artifact/org.apache.hadoop/hadoop-client>

Obrigada <3



THE
DEVELOPER'S
CONFERENCE



@kamilah_santos



in/kamila-santos-oliveira/



@kamila_code



@kamilahsantos





THE DEVELOPER'S CONFERENCE