Description: A Large Scale Kelkoo On line Advertising Data

Abstract: A novel, publicly available collection for recommendation systems that records the behavior of customers of the European leader in e Commerce advertising, Kelkoo, during month of June. This dataset describes the Germany. Complete dataset available at <URL of the website>

Dataset characteristics: Multi-Variate
Attribute Characteristics: Multi-Variate
Associated Tasks: Recommender Systems
Number of Instances:
Number of Attributes:
Missing Values: N/A
Area:
Date Donated:
Number of Web Hits:

Creators:

Massih R. Amini
Univ. Grenoble Alpes, CNRS/LIG
massih-reza.amini@imag.fr

Charlotte Laclau
Univ. Grenoble Alpes, CNRS/LIG
charlotte.laclau@univ-grenoble-alpes.fr}

Sumit Sidana
Univ. Grenoble Alpes, CNRS/LIG
sumit.sidana@imag.fr

Data Set Information:

We created this data by sampling and processing the www.kelkoo.com logs. The data records offers which were clicked (or shown) to the users of the www.kelkoo.com (and partners) in Germany as well as meta-information of these users and offers. The data contains 2 files training and testing set. Train set contains 291,485 users and associated 291,485 offers. Test set contains 278293 users and associated 380,803 offers. Both files contain additional attributes. Objective is to predict if a given user will click on a given offer.

Attribute Information:
Files, Instances and attributes
%--- Description Format
File Name
Number of Instances
Attribute: Type, Number of missing values (if any), Number of values
$---

1. train_de.csv.anon
Instances: 15,844,718
Attributes: 7
userid: Categorical, 291,485

offerid: Categorical, 2,158,859
countrycode: Categorical, 1
category: Integer, 271
merchant: Integer, 703
utcdate: Timestamp, **2016-06-01 02:00:17.0 to 2016-06-14 23:52:51.0**
rating: Binary, 0 or 1

2. test_de.csv.anon
Instances: 1,919,562
Attributes: 7
userid: Categorical, 278,293
offerid: Categorical, 380,803
countrycode: Categorical, 1
category: Integer, 267
merchant: Integer, 738
utcdate: Timestamp, **2016-06-14 23:52:51.0** to **2016-07-01 01:59:36.0**
rating: Binary, 0 or 1

<u>Relevant Papers</u>:

Sumit Sidana, Charlotte Laclau, Massih R. Amini, Gilles Vandelle, and André Bois-Crettez
KASANDR: A Large-Scale Dataset with Implicit Feedback for
Recommendation, SIGIR '17, August 07-11, 2017, Shinjuku, Tokyo, Japan