

**SISTEM REKOMENDASI E-COMMERCE AMAZON BERBASIS
COLLABORATIVE FILTERING**

SKRIPSI



oleh
DERRY SANTOSO
72140011

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
TAHUN 2018

**SISTEM REKOMENDASI E-COMMERCE AMAZON BERBASIS
COLLABORATIVE FILTERING**

SKRIPSI



© UKDW

Diajukan kepada Program Studi Sistem Informasi Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana
Sebagai Salah Satu Syarat dalam Memperoleh Gelar
Sarjana Komputer

Disusun oleh

DERRY SANTOSO
72140011

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
TAHUN 2018

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

Sistem Rekomendasi E-Commerce Amazon Berbasis Collaborative Filtering

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Sistem Informasi Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi keserjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa karya ilmiah ini adalah hasil plagiasi atau tiruan dari karya ilmiah lain, saya bersedia dikenai sanksi sesuai aturan yang berlaku di Universitas Kristen DutaWacana.

Yogyakarta, 26 Juni 2018



Derry Santoso

72140011

HALAMAN PERSETUJUAN

Judul Skripsi : Sistem Rekomendasi E-Commerce Amazon Berbasis Collaborative Filtering

Nama Mahasiswa : Derry Santoso

NIM : 72140011

Matakuliah : Skripsi

Kode : SI4426

Semester : Genap

Tahun Akademik : 2017/2018

Telah diperiksa dan disetujui di Yogyakarta,
Pada tanggal 26 Juni 2018

Dosen Pembimbing I



Yetli Oslan, SKom., M.T.

Dosen Pembimbing II



Erick Kurniawan, S.Kom., M.Kom.

HALAMAN PENGESAHAN

SISTEM REKOMENDASI E-COMMERCE AMAZON BERBASIS
COLLABORATIVE FILTERING

Oleh: DERRY SANTOSO / 72140011

Dipertahankan di depan Dewan Penguji Skripsi
Program Studi Sistem Informasi Fakultas Teknologi Informasi
Universitas Kristen DutaWacana - Yogyakarta
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer
pada tanggal
26 Juni 2018

Yogyakarta, 26 Juni 2018

Mengesahkan,

Dewan Penguji:

1. Yetli Oslan, S.Kom., M.T.
2. Erick Kurniawan, S.Kom., M.Kom.
3. Argo Wibowo, ST., MT.
4. Katon Wijana, S.Kom., M.T.



Dekan

(Budi Susanto, S.Kom, MT)

Ketua Program Studi

(Drs. Jong Jek Siang, M.Sc.)

ABSTRAKSI

Amazon adalah E-Commerce yang menjual berbagai macam barang, diantaranya: buku, komputer, dvd, dll. Dalam penelitian ini, data yang digunakan adalah data rating buku Amazon dari tahun 1995-2014. Penelitian dilakukan agar data rating tersebut dapat berguna untuk perkembangan E-Commerce dalam hal penjualan barang dengan membuat sistem rekomendasi dengan metode *Collaborative Filtering*. Sistem Rekomendasi akan memberikan rekomendasi sesuai dengan barang yang sudah pernah diberi rating oleh *user*.

Penelitian akan dimulai dari data rating yang masih dalam bentuk JSON. Data JSON akan dilakukan *filtering* berdasarkan *k-cores*. Data yang sudah difilter akan diolah menggunakan metode *Collaborative Filtering* dan akan dilakukan pembuatan model agar dapat dipanggil oleh API. API akan menyimpan model kedalam memori, lalu API akan diakses oleh aplikasi web. Hasil akhir penelitian adalah daftar rekomendasi yang sudah dicetak dalam aplikasi web.

Daftar rekomendasi yang dicetak ada 2 macam, yaitu: *item related* dan *user recommendation*. *Item Related* adalah daftar *item* yang berdekatan dengan *detail item* yang sedang dilihat oleh *user*. *User Recommendation* adalah rekomendasi yang akan didapatkan oleh user berdasarkan dengan rating yang sudah pernah diberikan oleh *user*.

Kata kunci: Collaborative Filtering, Machine Learning, Sistem Rekomendasi.

KATA PENGANTAR

Skripsi ini dicetak dengan tujuan memberikan informasi tentang bagaimana melakukan pengolahan data rating dan menggunakan data tersebut agar berguna untuk perkembangan sebuah E-Commerce. Penulis mengucapkan terima kasih kepada Tuhan Yang Maha Esa, pembimbing pertama ibu Yetli Oslan, pembimbing kedua pak Erick Kurniawan, dan teman-teman yang mendukung dalam penyusunan skripsi. Skripsi ini akan menjelaskan tentang data rating yang masih dalam bentuk JSON, dilakukan pengolahan, dan dicetak pada aplikasi berbasis web sebagai sistem rekomendasi.

Jika ada kata-kata yang salah penulis mohon maaf sebesar-besarnya. Atas perhatian saudara penulis mengucapkan terima kasih.

Yogyakarta, Juni 2018

Penulis Skripsi,



Derry Santoso

Daftar Isi

PERNYATAAN KEASLIAN SKRIPSI	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PENGESAHAN	v
ABSTRAKSI	vi
KATA PENGANTAR	vii
BAB 1	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	1
1.3. Batasan Masalah	2
1.4. Spesifikasi Sistem.....	2
1.4.1. Spesifikasi software	2
1.4.2. Spesifikasi hardware	2
1.4.3. Spesifikasi aplikasi	3
1.4.4. Spesifikasi kecerdasan pembangun.....	3
1.4.5. Spesifikasi kecerdasan pengguna aplikasi	3
1.5. Tujuan dan Manfaat Penelitian	3
1.6. Metodologi Penelitian.....	3
1.7. Sistematika Penulisan	4
BAB 2	5
2.1. <i>Machine Learning</i>	5
2.2. <i>Collaborative Filtering</i>	5
2.2.1. <i>Collaborative Filtering Cost Function</i>	6
2.2.2. <i>Collaborative Filtering Stochastic Gradient Descent</i>	7
2.2.3. <i>Regularization</i>	8
2.3. <i>Cross Validation</i>	9
2.4. <i>Cosine Similarity</i>	10
BAB 3	11
3.1. Data Penelitian.....	11
3.2. Gambaran Flowchart Penelitian.....	11
3.2. Gambaran Sistem.....	16
3.2.1 <i>DFD</i>	16
3.2.2 <i>Use Case</i>	17
3.2.3 <i>Activity Diagram</i>	20
BAB 4	21

4.1 Pembuatan Sistem.....	21
4.1.1 Filter Amazon Rating JSON Menggunakan MongoDB	21
4.1.2 Pembuatan Model dengan Metode Collaborative Filtering	22
4.1.3 Menyiapkan Model untuk produksi dan Flask API	28
4.1.4 Aplikasi Web Laravel	31
4.2 Analisis	36
BAB 5	39
5.1 Kesimpulan	39
5.2 Saran	39
Daftar Pustaka.....	40
Lampiran.....	44

©UKDW

ABSTRAKSI

Amazon adalah E-Commerce yang menjual berbagai macam barang, diantaranya: buku, komputer, dvd, dll. Dalam penelitian ini, data yang digunakan adalah data rating buku Amazon dari tahun 1995-2014. Penelitian dilakukan agar data rating tersebut dapat berguna untuk perkembangan E-Commerce dalam hal penjualan barang dengan membuat sistem rekomendasi dengan metode *Collaborative Filtering*. Sistem Rekomendasi akan memberikan rekomendasi sesuai dengan barang yang sudah pernah diberi rating oleh *user*.

Penelitian akan dimulai dari data rating yang masih dalam bentuk JSON. Data JSON akan dilakukan *filtering* berdasarkan *k-cores*. Data yang sudah difilter akan diolah menggunakan metode *Collaborative Filtering* dan akan dilakukan pembuatan model agar dapat dipanggil oleh API. API akan menyimpan model ke dalam memori, lalu API akan diakses oleh aplikasi web. Hasil akhir penelitian adalah daftar rekomendasi yang sudah dicetak dalam aplikasi web.

Daftar rekomendasi yang dicetak ada 2 macam, yaitu: *item related* dan *user recommendation*. *Item Related* adalah daftar *item* yang berdekatan dengan *detail item* yang sedang dilihat oleh *user*. *User Recommendation* adalah rekomendasi yang akan didapatkan oleh user berdasarkan dengan rating yang sudah pernah diberikan oleh *user*.

Kata kunci: Collaborative Filtering, Machine Learning, Sistem Rekomendasi.

BAB 1

PENDAHULUAN

1.1. Latar Belakang

(O'Brien & Marakas, 2007) E-Commerce bukan hanya sekedar tempat untuk melakukan jual beli, tetapi juga tempat untuk berkembang, promosi, mengirim, dan servis produk. Kelebihan E-commerce adalah sebesar/sekecil apapun bisnis retailnya, jarak fisik tidak akan menjadi penghalang untuk melakukan bisnis. Semua pebisnis dan konsumen di bumi memiliki peluang untuk melakukan transaksi pada E-commerce.

Aplikasi E-Commerce yang akan dibuat adalah C2C (Customer to Customer). Aplikasi diharapkan dapat memberikan rekomendasi yang bersifat personal untuk setiap pembeli. Sistem Rekomendasi dengan Metode Collaborative Filtering akan memberikan rekomendasi berdasarkan : sejarah produk yang sudah diberi rating oleh pembeli terdahulu, sejarah *user* yang sudah memberi rating pada produk, fitur referensi tiap *user* terhadap produk, sejarah rating pembelian *user*.

Penelitian akan berfokus pada bagaimana pengukuran keakuratan rekomendasi di dalam aplikasi E-Commerce yang sudah siap produksi. Pembuatan model Sistem Rekomendasi akan dilakukan menggunakan bahasa pemrograman python. Model yang sudah jadi akan dikirimkan melalui Application Programming Interface (API) python ke Laravel. Framework Laravel akan digunakan sebagai dasar aplikasi E-Commerce berbasis Web.

1.2. Rumusan Masalah

1. Bagaimana mengimplementasikan Sistem Rekomendasi pada E-Commerce dalam memberikan rekomendasi buku yang bersifat pribadi untuk setiap pelanggan?
2. Bagaimana pengukuran keakuratan sistem rekomendasi yang akan diimplementasikan ke dalam Aplikasi E-Commerce berbasis Web?

1.3. Batasan Masalah

1. Data *User* Amazon tahun 1995 – 2014 yang sudah pernah memberi rating pada produk buku dan kode produk buku yang diberi rating
2. Metadata produk Amazon dalam kategori buku pada dataset tahun 1995 – 2014.
3. Semua data berasal dari repositori Julian McAuley institusi UCSD (University of California San Diego) (R. He, 2016).
4. Sample rating akan diambil dari setiap *user* sebanyak 10 rating sebagai test set, sisa rating yang tidak diambil akan digunakan sebagai training set.
5. Stakeholder yang terkait adalah pembeli (*user*) dalam aplikasi E-Commerce berbasis web.
6. Metode machine learning yang digunakan adalah Collaborative Filtering.
7. Bahasa yang digunakan untuk memproses data dan pembuatan model sistem rekomendasi adalah Python dan Flask untuk memproses API.
8. Framework yang digunakan sebagai dasar aplikasi E-Commerce adalah Laravel dengan bahasa pemrograman PHP.
9. Aplikasi E-Commerce web akan berfokus pada fitur yang dibutuhkan untuk melakukan penelitian sistem rekomendasi.

1.4. Spesifikasi Sistem

1.4.1. Spesifikasi software

1. Sistem Operasi Windows 10 64 bit
2. Browser Mozilla Firefox / Chrome versi terbaru

1.4.2. Spesifikasi hardware

1. 3.5 Ghz i7 6700HQ
2. 16 GB RAM
3. 1 Terabyte HDD
4. DirectX 12 graphics device
5. Nvidia GTX 1060 6GB VRAM
6. Akses Internet

7. Monitor 15", resolusi (1920x1080)
8. Keyboard dan Mouse

1.4.3. Spesifikasi aplikasi

1. Sistem Rekomendasi mampu memberikan rekomendasi yang bersifat personal kepada *user* dengan metode Collaborative Filtering
2. Model Sistem Rekomendasi yang digunakan dalam aplikasi berbasis web sudah teroptimisasi dan siap untuk tahap produksi

1.4.4. Spesifikasi kecerdasan pembangun

1. Kemampuan dalam pengolahan data menggunakan Python.
2. Kemampuan pembuatan aplikasi web menggunakan framework Laravel dengan bahasa pemrograman PHP.
3. Kemampuan melakukan tes keakuratan dan optimisasi model sistem rekomendasi.

1.4.5. Spesifikasi kecerdasan pengguna aplikasi

1. Mampu menggunakan komputer dan mengakses website melalui browser
2. Memahami istilah bahasa Inggris yang digunakan dalam aplikasi

1.5. Tujuan dan Manfaat Penelitian

1. Membuat suatu sistem rekomendasi yang dapat memberikan rekomendasi buku ke tiap pelanggan secara pribadi menggunakan metode Collaborative Filtering.
2. Membuat Sistem Rekomendasi yang sudah siap digunakan sampai ke tahap produksi untuk aplikasi berbasis web.

1.6. Metodologi Penelitian

1. Studi Pustaka

Mempelajari bagaimana cara pembuatan sistem rekomendasi menggunakan metode Collaborative Filtering melalui website dan jurnal ilmiah. Lalu mempelajari cara pemrosesan data menggunakan python dan cara mengirim

model yang sudah jadi ke aplikasi web berbasis Laravel menggunakan API. Terakhir adalah mempelajari cara optimisasi sistem rekomendasi pada aplikasi web dan cara pengukuran keakuratan model yang dibuat.

2. Pengumpulan Data

Mengumpulkan data review buku Amazon melalui repositori Julian McAuley institusi UCSD

3. Analisis Data

Analisa menggunakan sample rating yang diambil dari setiap *user* sebanyak 10 rating sebagai test set, sisa rating yang tidak diambil akan digunakan sebagai training set.

4. Rancangan Sistem

Mendesain sistem yang akan dibangun

5. Implementasi Sistem

Pembuatan model sistem rekomendasi menggunakan metode Collaborative Filtering dengan Python, pembuatan aplikasi E-Commerce Web menggunakan Laravel, membuat API untuk mengirim data dari python ke Laravel, tampilkan data hasil rekomendasi melalui aplikasi web.

6. Pengujian dan Analisis Sistem

Melakukan uji keakuratan sistem rekomendasi menggunakan metode cross validation pada 10 data rating dari setiap *user* sebagai test set.

7. Penyelesaian Laporan

Membuat kesimpulan, saran, kelebihan dan kelemahan sistem yang dibangun.

1.7 Sistematika Penulisan

Bab 1 akan menjelaskan tentang latar belakang, rumusan masalah, batasan masalah, spesifikasi sistem, tujuan dan manfaat penelitian, metodologi penelitian, dan sistematika penulisan. Bab 2 berisi penjelasan teori metode yang akan digunakan dalam penelitian. Bab 3 berisi rancangan sistem, sedangkan Bab 4 akan dilakukan pembuatan, penerapan, dan analisis sistem. Bab 5 berisi penutup yang berisi tentang kesimpulan dan saran penelitian.

BAB 5 PENUTUP

5.1 Kesimpulan

Hasil penelitian sistem rekomendasi dengan metode Collaborative Filtering dapat mengeluarkan hasil akhir RMSE yang cukup baik sebesar 0.872765146. Tetapi walaupun hasil RMSE cukup baik, rekomendasi visual yang keluar dalam aplikasi web cukup sulit untuk dinilai secara intuisi apakah hasil *related item* pada halaman detail *item* dan *user recommendation* pada halaman *home* benar-benar cocok untuk *user*.

5.2 Saran

- a. Menggunakan Metode Content Filtering sebagai pelengkap Collaborative Filtering agar hasil rekomendasi tidak sepenuhnya bergantung dengan dataset, daftar rekomendasi visual dalam web lebih masuk akal, dan rekomendasi dapat dinilai secara intuisi.
- b. Menggunakan Metode Content Filtering untuk memberikan rekomendasi *related item* pada *item* yang belum pernah diberi rating oleh *user*.
- c. Penelitian lebih lanjut tentang metode untuk memproses item dan rating baru yang tidak ada dalam dataset. Item dan rating baru dapat digunakan untuk melatih ulang model agar performa model meningkat (RMSE bertambah kecil).
- d. Penelitian lebih lanjut untuk memfilter *user* yang sejarah pembeliannya terlalu bervariasi agar tidak menjadi *noise* (contoh: *user* yang membeli buku komputer juga membeli buku drama dan fiksi).

Daftar Pustaka

- Barwicki, A. (2017, June 16). *Recommendation System: Simple Theory of Collaborative Filtering*. Retrieved from Adrian's Space:
<http://adrianbarwicki.com/2017/06/16/collaborative-filtering-recommendation-systems-simple-theory/>
- Bluma, A. L., & Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 245-271.
- Bossenbroek, H., & Gringhuis, H. (2014). Recommendation in E-Commerce. *Luminis Recommendation Services*.
- Isinkayea, F., Folajimib, Y., & Ojokohc, B. (2015). Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, 261–273.
- Joonseok Lee, M. S. (2012). A Comparative Study of Collaborative Filtering. *arXiv:1205.3193v1 [cs.IR]*.
- Kohavi, R., & Provost, F. (1998). Glossary of Terms. *Machine Learning*, 271-274.
- Netflix. (2009). *Netflix Prize Leaderboard*. Retrieved from Netflix:
<https://www.netflixprize.com/leaderboard.html>
- Ng, A. (2017). *Machine Learning by Stanford University*. Retrieved from Coursera:
<https://www.coursera.org/learn/machine-learning/home/welcome>
- O'Brien, J., & Marakas, G. (2007). *Management Information System 10th ed*. New York: McGraw Hill.
- Perone, C. S. (2013, 9 12). *Machine Learning :: Cosine Similarity for Vector Space Models (Part III)*. Retrieved from Terra Incognita:
<http://blog.christianperone.com/2013/09/machine-learning-cosine-similarity-for-vector-space-models-part-iii/>
- R. He, J. M. (2016). Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. *WWW*.
- Rosenthal, E. (2016, March 16). *Explicit Matrix Factorization: ALS, SGD, and All That Jazz*. Retrieved from Medium Insight Data Science:
<https://blog.insightdatascience.com/explicit-matrix-factorization-als-sgd-and-all-that-jazz-b00e4d9b21ea>
- Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2002). Recommender systems for large-scale e-commerce: Scalable neighborhood formation using clustering. *Proceedings of the fifth international conference on computer and information technology* (pp. 1-6). Citeseer.