



UNIVERSITAS INDONESIA

**Implementasi Teknik *Information Retrieval* untuk Fitur
Rekomendasi Produk pada Sistem *e-Commerce* PT. Kudo Teknologi
Indonesia**

LAPORAN KERJA PRAKTIK

Haryo Akbarianto Wibowo

1306403850

**FAKULTAS ILMU KOMPUTER
PROGRAM STUDI Ilmu Komputer**

DEPOK

Agustus 2016

Universitas Indonesia

HALAMAN PERSETUJUAN DOSEN MATA KULIAH KERJA PRAKTIK

Laporan ini diajukan oleh :
Nama : Haryo Akbarianto Wibowo
NPM : 1306403850
Program Studi : Ilmu Komputer
Judul Kerja Praktik : Implementasi Teknik Information Retrieval untuk
Fitur Rekomendasi Produk pada Sistem e-
Commerce PT. Kudo Teknologi Indonesia

Telah berhasil diselesaikan laporan kerja praktik untuk fakultas dan dipresentasikan hasil kerja praktiknya dalam forum seminar kerja praktik sebagai persyaratan yang harus dipenuhi dalam mata kuliah Kerja Praktik.

DOSEN MATA KULIAH KERJA PRAKTIK,

(Rahmad Mahendra, S.Kom., M.Sc.)

Ditetapkan di : Depok

Tanggal : X

Universitas Indonesia

ABSTRAK

Laporan ini berisi tentang hasil kerja praktik yang dilaksanakan pada tanggal 13 Juni 2016 sampai 19 Agustus 2016 di PT. Kudo Teknologi Indonesia. Penulis ditetapkan sebagai *Data Engineer* pada PT. Kudo Teknologi Indonesia. Pekerjaan yang didapatkan oleh penulis adalah perancangan implementasi teknik *Information Retrieval* untuk fitur rekomendasi produk pada PT. Kudo Teknologi Indonesia. Pada saat menjalani pekerjaan tersebut, penulis mendapat tugas untuk mencari rekomendasi barang yang mirip dan komplemen dengan barang yang ditargetkan. Beberapa teknologi penulis gunakan untuk melaksanakan pekerjaan tersebut. Pada laporan ini terdapat analisis perbedaan kerangka acuan kerja praktik dengan pelaksanaan kerja praktik. Kerja praktik yang penulis lakukan mempunyai relevansi terhadap mata kuliah yang diajarkan di Fakultas Ilmu Komputer.

Kata kunci: *data engineer*, *recommendation engine*, *information retrieval*, kerja praktik

DAFTAR ISI

HALAMAN PERSETUJUAN DOSEN MATA KULIAH KERJA PRAKTIK	10
ABSTRAK.....	11
DAFTAR ISI.....	12
DAFTAR GAMBAR	14
DAFTAR TABEL.....	15
DAFTAR LAMPIRAN.....	16
BAB 1 PENDAHULUAN	17
1.1. Proses Pencarian Kerja Praktik	17
1.2. Tempat Kerja Praktik	19
1.2.1. Profil Tempat Kerja Praktik.....	19
1.2.2. Posisi Penempatan Pelaksana Kerja Praktik dalam Struktur Organisasi	
20	
BAB 2 ISI	22
2.1. Pekerjaan dalam Kerja Praktik	22
2.1.1. Latar Belakang Pekerjaan	22
2.1.2. Tinjauan Pustaka	23
2.1.2.1. Information Retrieval	23

	13
2.1.2.2. <i>Recommender System</i>	24
2.1.2.3. <i>Vector Space Model</i>	24
2.1.2.4. Bobot Tf-Idf	25
2.1.2.5. Fungsi Kemiripan	25
2.1.2.6. RESTful API	27
2.1.2.7. <i>Complementary Good</i>	27
2.1.3. Metodologi	27
2.1.4. Teknologi	28
2.1.5. Gambaran Pekerjaan	30
2.2. Analisis	35
2.2.1. Pelaksanaan Kerja Praktik	35
2.2.2. Relevansi dengan Perkuliahan di Fasilkom UI	37
2.2.3. Pembelajaran <i>Soft Skill</i> dari Kerja Praktik	38
BAB 3 PENUTUP	39
3.1. Kesimpulan	39
3.2. Saran	39
DAFTAR REFERENSI	41
LAMPIRAN 1 KERANGKA ACUAN KERJA PRAKTIK	44
LAMPIRAN 2 LOG KERJA PRAKTIK	45

DAFTAR GAMBAR

Gambar 1 Posisi penulis pada pelaksanaan kerja praktik.....	21
Gambar 2 Visualisasi Vector Space Model.....	24
Gambar 3 Alur gambaran kerja penulis.....	30
Gambar 4 Pemasangan kategori komplemen.....	34

DAFTAR TABEL

[Daftar tabel, ditulis serupa dengan daftar isi]

DAFTAR LAMPIRAN

[Daftar Lampiran, ditulis serupa dengan daftar isi]

BAB 1

PENDAHULUAN

1.1. Proses Pencarian Kerja Praktik

Proses penulis mencari tempat kerja praktik lumayan panjang. Penulis memulai pencarian lowongan tempat magang sejak awal bulan Maret 2016. Penulis lebih fokus terhadap pencarian lowongan pekerjaan pada perusahaan *startup* dikarenakan penulis mendengar bahwa di perusahaan *startup* peluang belajarnya lebih besar dibanding yang bukan perusahaan *startup*. beberapa perusahaan yang penulis coba untuk daftar yaitu Tokopedia, Traveloka, Bukalapak, Kudo, Suitmedia, dan Indivara. Dari 6 perusahaan tersebut, hanya Indivara yang bersedia untuk menerima penulis. Karena PT. Kudo menawarkan peluang lagi untuk mendaftar, penulis memilih untuk masuk PT. Kudo dikarenakan PT. Kudo merupakan perusahaan *startup* dan bidang yang ditawarkan sesuai dengan minat penulis, yaitu bidang data.

Pada awalnya, penulis mendaftar pada perusahaan Tokopedia, Bukalapak, dan Traveloka yang membuka lowongan pekerjaan magang. Penulis mendapatkan kontak *e-mail* dari *website* masing-masing perusahaan tersebut. Setelah penulis mendapatkan kontak *e-mail*, penulis mengirimkan CV dan permohonan untuk melamar magang di perusahaan tersebut. Adapun posisi yang penulis minta pada perusahaan Tokopedia dan Bukalapak yaitu *Software Engineer*. Sedangkan posisi yang penulis minta pada perusahaan Traveloka yaitu *Data Analyst*.

Respon tercepat pertama yang penulis dapatkan yaitu dari perusahaan Tokopedia. Pertama-tama penulis diberikan soal tentang *programming* secara *online*. penulisngnya, karena ketidaktelitian penulis, penulis gagal pada tahap ini.

Respon tercepat kedua didapatkan dari perusahaan Traveloka. Perusahaan tersebut mengundang penulis untuk melakukan wawancara di kantornya. Wawancara yang dilakukan oleh perusahaan Traveloka adalah wawancara dalam menjawab soal. Soal yang diberikan adalah soal tentang *binary search tree* dan basis data. Penulis dapat menyelesaikan masalah tersebut dengan baik. Setelah penulis selesai melakukan wawancara, penulis menunggu hasil dari wawancara tersebut. Sayangnya, perusahaan ini menolak penulis dengan *e-mail* yang dikirimkan sebulan setelah wawancara tersebut.

Sudah dua bulan perusahaan Bukalapak tidak merespon lamaran yang penulis kirimkan. Penulis menganggap bahwa perusahaan Bukalapak menolak penulis. Penulis mencari informasi kembali mengenai kerja praktik. Penulis mendapat kabar bahwa perusahaan Kudo dan Suitmedia membuka lowongan magang. Perusahaan Kudo menawarkan posisi sebagai *Data Engineer* dan Suitmedia menawarkan posisi sebagai *Software Engineer*. Penulis mendaftar magang ke kedua perusahaan tersebut. Penulis menjalani tes wawancara kedua perusahaan tersebut dalam waktu yang berdekatan. Namun, penulis ditolak oleh keduanya.

Setelah penulis ditolak kedua perusahaan tersebut, penulis mendaftar salah satu jasa penyedia lowongan magang yaitu Quint. Setelah beberapa hari, penulis mendapat tawaran tes wawancara untuk magang di salah satu perusahaan yang bekerjasama dengan Quint. Perusahaan itu adalah PT. Indivara Group. Tawaran ini muncul pada tengah bulan Mei 2016.

Pada saat tawaran tes wawancara untuk magang tersebut ditawarkan, penulis mengikuti lomba Hackaton yang ditawarkan oleh perusahaan Kudo. Lomba tersebut diadakan pada tanggal 28 Mei 2016 sampai 29 Mei 2016. Pada saat lomba tersebut, penulis menanyakan tentang tawaran magang di Kudo ke CEO Kudo (*Chief*

Executive Officer). Beliau berkata bahwa jika penulis ikut lomba ini, penulis bisa diterima tanpa tes. Penulis perlu mengirimkan *e-mail* ke HRD (*Human Research Development*) perusahaan Kudo tentang lamaran magang terlebih dahulu.

Dua hari setelah lomba tersebut, penulis melakukan tes wawancara terhadap perusahaan PT. Indivara Group. Posisi yang ditawarkan adalah *Software Engineer*. Setelah melakukan tes wawancara, penulis diterima magang di perusahaan ini. Penulis meminta ke HRD perusahaan ini untuk memberikan waktu seminggu karena penulis ingin mencoba mendaftar ke perusahaan Kudo lagi. Permintaan penulis diterima oleh HRDnya.

Penulis mencoba kembali mendaftar Kudo dengan mengirimkan langsung *e-mail* ke CEOnya. Dalam beberapa menit, penulis mendapat respon bahwa penulis diterima di perusahaan Kudo sebagai *Data Engineer*. Penulis mendapatkan posisi ini karena penulis menulis posisi yang diminati pada kiriman *e-mail* tersebut. Setelah itu, penulis memberitahu perusahaan PT. Indivara Group bahwa penulis diterima di perusahaan Kudo.

Penulis lebih memilih perusahaan Kudo disebabkan karena minat penulis adalah bidang data sesuai yang ditawarkan oleh perusahaan Kudo. Suasana pekerjaan yang terbuka dan jam waktu yang fleksibel membuat penulis memilih perusahaan Kudo sebagai tempat magang.

1.2. Tempat Kerja Praktik

1.2.1. Profil Tempat Kerja Praktik

Kudo adalah perusahaan *startup* yang bergerak sebagai *platform* yang memberikan peluang usaha menjadi *reseller* produk dari berbagai *e-commerce*. Perusahaan ini didirikan pada Juli 2014. Startup ini mengembangkan solusi praktis untuk

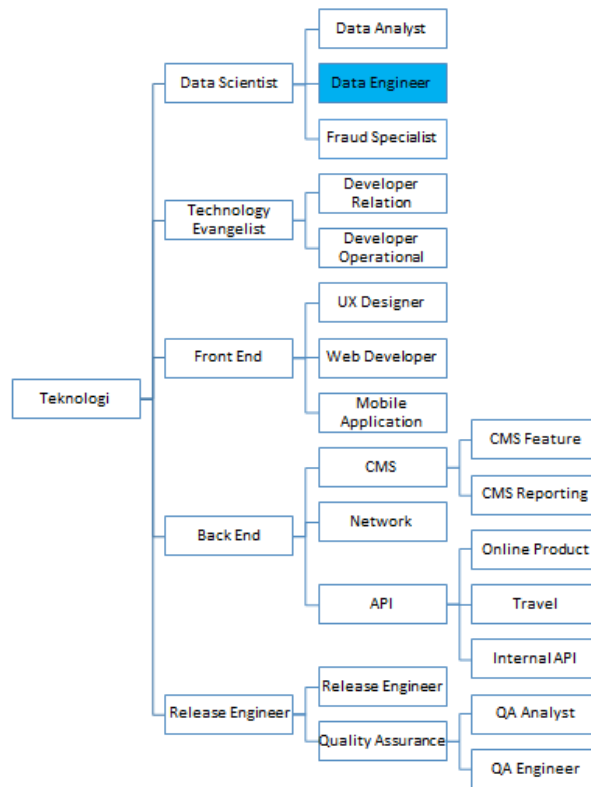
marketplace dan ekosistem pembayaran di Indonesia. Perusahaan Kudo memberikan solusi untuk masyarakat Indonesia yang tidak memiliki akses dan tidak mau bertransaksi *online* untuk melakukan pembelian *online shop* secara tunai. Perusahaan Kudo bekerja sama dengan perusahaan *e-commerce* lainnya untuk mengembangkan solusi tersebut [1].

Kudo membuat aplikasi *mobile* dan website untuk menjalankan bisnis utamanya. Pengguna aplikasi tersebut disebut agen. Agen akan menjual barang dagangan yang dipesan secara *online* dari aplikasi Kudo kepada konsumennya.

Secara organisasi, perusahaan Kudo memiliki dua bidang utama yaitu bidang bisnis yang dibawah oleh *Chief Executive Officer* (COO) dan bidang teknologi yang dikepalai oleh *Chief Technology Officer* (CTO) yang juga bertindak sebagai *Chief Executive Officer* (CEO). Bidang bisnis fokus terhadap kegiatan bisnis seperti *marketing*, *product*, *accounting*, dan *marketing*. Sedangkan bidang teknologi lebih berfokus terhadap pengembangan teknologi yang dipasarkan. Dalam masing-masing bidang terdapat posisi yang diketuai oleh seorang *Head Division*. Dibawah *Head Division* masih terdapat *Lead Position* yang bertugas untuk membantu *Head Division* mengawasi bagian bidang yang lebih spesifik.

1.2.2. Posisi Penempatan Pelaksana Kerja Praktik dalam Struktur Organisasi

Pada saat menjalankan kerja praktik, penulis ditempatkan sebagai *Data Engineer* yang berada di bawah divisi data. *Data Engineer* berperan untuk mengolah data yang terkumpul di perusahaan Kudo. Dari data yang terkumpul tersebut dicari abstraksi model dengan menggunakan algoritma yang berhubungan dengan data dan menggunakannya untuk memecahkan masalah yang ada. Gambar 2 menjelaskan struktur organisasi yang memuat penempatan penulis.



Gambar 1 Posisi penulis pada pelaksanaan kerja praktik

BAB 2

ISI

2.1. Pekerjaan dalam Kerja Praktik

2.1.1. Latar Belakang Pekerjaan

Perusahaan Kudo merupakan perusahaan *startup* dibidang *e-commerce* yang didalamnya banyak melakukan transaksi antar penggunanya. Transaksi antar pengguna ini disimpan didalam suatu basis data dimana jumlah dari datanya lumayan besar. Data yang besar ini mempunyai potensi untuk diolah untuk memecahkan suatu permasalahan atau meningkatkan fitur yang ada pada perusahaan Kudo untuk memperbesar keuntungan penjualannya. Untuk memecahkan masalahnya, perlu dibuat suatu abstraksi model yang akan menjadi barang inti untuk memecahkan masalah tersebut. Pembuatan abstraksi model inilah yang menjadi pekerjaan penulis.

Pekerjaan yang diberikan untuk penulis dan rekan penulis yang berada pada posisi *Data Engineer* yaitu membuat dua fitur. Dua fitur itu adalah *recommender system* dan *agent profiling*. Kedua fitur tersebut belum pernah dikembangkan oleh perusahaan ini. Penulis diperintahkan untuk mengerjakan fitur *recommender system* dalam melakukan kerja praktik disini.

Dalam melakukan pekerjaannya, penulis tidak bekerja secara sendiri. Penulis dibantu oleh rekan penulis yang bernama Hayyu. Dia merupakan mahasiswa Institut Teknologi Bandung (ITB) yang juga menjadi bagian dari anggota magang di perusahaan Kudo. Rekan penulis hanya melakukan magang selama 8 minggu saja, sehingga penulis mengerjakan fitur ini sendiri.

Fitur *recommender system* dibuat karena banyaknya barang yang tersedia untuk dijual kembali oleh agen. Fitur ini berguna untuk memudahkan agen mencari barang yang mirip yang sesuai dengan barang yang diminatinya. *recommender system* yang akan dibuat dibagi menjadi dua, yaitu *recommender system* berdasarkan personalisasi dan *recommender system* berdasarkan kemiripan barang. Penulis ditugaskan untuk membuat *recommender system* berdasarkan kemiripan barang.

Recommender system yang penulis buat juga dibagi menjadi dua, yaitu *recommender system* untuk pencarian rekomendasi barang yang mirip dan rekomendasi barang yang komplemen terhadap barang yang dilihat oleh pengguna..

Fitur *recommender system* tersebut juga diimplementasikan kedalam suatu API. Pembuatan API ini juga menjadi tugas penulis. API ini berguna untuk memberikan akses kepada divisi lain untuk mencoba fitur *recommender system* ini.

Penulis juga diminta untuk membuat visualisasi terhadap fitur *recommender system* ini melalui tampilan. Mentor kami menyarankan untuk membuat website sederhana untuk visualisasinya. Website ini nantinya akan diperlihatkan pada saat presentasi ke semua divisi untuk didemokan.

2.1.2. Tinjauan Pustaka

2.1.2.1. Information Retrieval

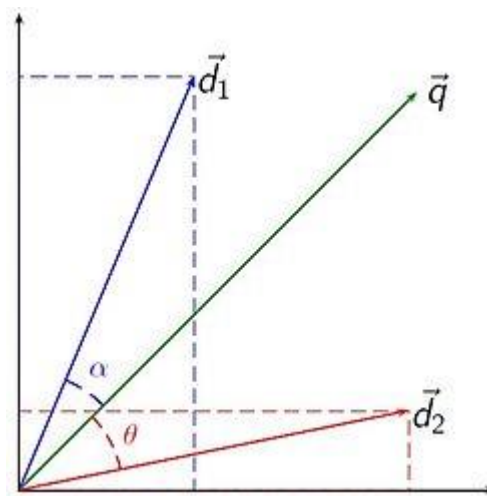
Information Retrieval (IR) adalah ilmu yang mencari informasi yang dibutuhkan dari suatu materi yang bentuknya tidak terstruktur yang menjawab suatu *query* [21]. Tujuan utama dari IR adalah untuk mencari informasi yang relevan dari *query* yang diberikan terhadap kumpulan dokumen yang tersedia [2] .

2.1.2.2. *Recommender System*

Recommender System adalah alat perangkat lunak yang memberikan rekomendasi atau saran sesuatu terhadap penggunaanya. Rekomendasi tersebut berhubungan dengan kegiatan memilih sesuatu yang dilakukan oleh pengguna. Contohnya yaitu memilih barang yang akan dibeli dan memilih film yang akan ditonton [3]. Tujuan utama dari Recommender System adalah untuk membuat rekomendasi yang bermakna kepada pengguna untuk barang atau produk yang bisa menarik perhatian mereka. [4]

2.1.2.3. *Vector Space Model*

Vector Space Model (VSM) adalah teknik untuk merepresentasikan suatu dokumen yang tersedia pada koleksi menjadi titik pada space. Titik tersebut adalah vektor pada suatu ruang vektor. Jarak titik satu sama lain menyatakan kedekatan dan kemiripan [5].



Gambar 2 Visualisasi *Vector Space Model*

2.1.2.4. Bobot Tf-Idf

Term Frequency - Inverse Document Frequency (Tf-idf) adalah salah satu cara untuk memberi pembobotan untuk representasi titik pada VSM [6][7]. Pembobotan tf-idf dirumuskan sebagai berikut:

$$w_{t,d} = tf_{t,d} * \log \frac{D}{f_{t,D}}$$

Dimana:

- $tf_{t,d}$ adalah frekuensi kata t muncul pada dokumen d
- D adalah jumlah dokumen yang ada di koleksi
- $f_{t,D}$ adalah frekuensi jumlah dokumen yang muncul pada koleksi dimana dokumen tersebut memiliki kata t .

2.1.2.5. Fungsi Kemiripan

Untuk mengukur kedekatan antar dokumen yang direpresentasikan dari VSM, berikut adalah fungsi kemiripan yang digunakan untuk melihat kedekatannya:

1. *Cosine Similarity*

Cosine similarity adalah salah satu pengukuran kedekatan antar titik pada ruang vektor [8]. *Cosine similarity* juga bisa diaplikasikan untuk mencari kedekatan ruang titik pada VSM. Rumus *cosine similarity* didefinisikan sebagai berikut:

$$sim(a, b) = \frac{a.b}{\sqrt{(a.a)(b.b)}}$$

Dimana:

- a dan b adalah vektor suatu titik yang merepresentasikan suatu dokumen yatau query yang direpresentasikan oleh suatu bobot. a dan b tidak negatif

2. Jaccard Similarity

Jaccard Similarity adalah salah satu pengukuran kedekatan antar titik pada *keyword* yang dimasukkan dengan dokumen yang tersedia. [9]. Bobot yang diberikan untuk semua kata pada setiap dokumen harus berupa biner (0 atau 1). Bobot akan bernilai 0 jika kata tersebut tidak ada pada dokumen dan 1 jika sebaliknya. Berikut adalah rumus dari *Jaccard Similarity*:

$$sim(a, b) = \frac{a \cap b}{a \cup b}$$

Dimana:

- a dan b adalah himpunan kata yang berada pada dokumen a dan b

3. Euclidean Distance

Euclidean Distance adalah salah satu pengukuran kedekatan antar titik pada ruang vektor yang bisa digunakan untuk VSM. *Euclidean Distance* melihat jarak antara kedua titik dilihat dari panjang garis lurus yang digambarkan untuk kedua titik tersebut [10]. Berikut adalah rumus dari *Euclidean Distance*:

$$sim(a, b) = \sqrt{\sum_{i=0}^n a_i^2 - b_i^2}$$

Dimana:

- a_i, b_i adalah nilai vektor a dan b pada dimensi ke-i.

4. *Chevychev Distance*

Chevychev Distance adalah salah satu cara untuk mencari kedekatan antar titik pada ruang vektor. *Chevychev Distance* menghitung jumlah perbedaan kedua vektor yang terbesar pada setiap dimensi [11]. Berikut adalah rumus dari *Chevychev Distance*:

$$sim(a, b) = \max_i(a_i - b_i)$$

Dimana:

- a_i, b_i adalah nilai vektor a dan b pada dimensi ke-i.

2.1.2.6. RESTful API

Representational State Transfer Application Program Interface (RESTful API) adalah aplikasi program yang menggunakan HTTP Request untuk GET, PUT, POST, dan DELETE suatu data. RESTful API memecah transaksi yang masuk menjadi beberapa modul yang kecil. Modul yang kecil tersebut akan menangani bagian tertentu dari transaksi ini [12].

2.1.2.7. Complementary Good

Complementary Good adalah barang yang digunakan untuk dihubungkan dengan barang lain dimana *Complementary Good* tersebut akan menghasilkan *demand* terhadap barang yang berelasi dengan barang tersebut [13]. Barang komplemen juga merupakan barang yang biasanya dibeli bersamaan [20]. Contoh dari *Complementary Good* adalah laptop dengan tetikus.

2.1.3. Metodologi

Dalam pelaksanaan kerja praktik pada tim data di Kudo, penulis tidak menggunakan metode pengembangan secara khusus. Penulis diberikan kebebasan mengenai *timeline* pengerjaan fitur *recommender system*. Walaupun diberi kebebasan, penulis harus memperhatikan *deadline* untuk presentasi progress fitur tersebut ke Kudo.

Selama menjalankan pelaksanaan kerja praktik, penulis sering berkomunikasi dengan mentor mengenai *requirement* dan kemajuan dari fitur *recommender system* ini. Komunikasi ini berlangsung bebas sesuai dengan kesepakatan penulis dan mentor. Penulis juga mengikuti budaya *weekly meeting* khusus untuk *technology team* yang dilakukan seminggu sekali untuk membahas kemajuan dari proyek yang dikerjakan oleh setiap divisi. Selain *weekly meeting*, penulis juga mengikuti *stand up meeting* yang membahas kemajuan untuk tiap proyek yang sedang dikerjakan pada tim data. *stand up meeting* dilakukan dalam waktu tidak tentu sesuai dengan keinginan *head* dari tim data.

Adapun alur pekerjaan penulis dalam melaksanakan kerja praktik di Kudo adalah sebagai berikut:

1. Riset tentang bagaimana cara pencarian barang yang mirip
2. Pembuatan program dan website untuk uji fungsi kemiripan yang ditemui dari hasil riset
3. Uji fungsi kemiripan ke beberapa orang untuk dicari mana fungsi kemiripan yang terbaik
4. Buat modul dan RESTful-API dari sistem rekomendasi berdasarkan barang yang mirip
5. Riset tentang barang komplementer
6. Buat modul dan RESTful-API untuk barang komplementer
7. Buat website untuk demo presentasi akhir
8. Presentasi akhir

2.1.4. Teknologi

Berikut adalah teknologi yang penulis pakai saat menjalankan Kerja Praktik:

1. Bitbucket

Bitbucket adalah *web-based hosting service* untuk pekerjaan yang menggunakan Git atau Mercurial untuk melakukan kontrol pada revisi [14]. Bitbucket dimiliki oleh perusahaan Atlassian. Beberapa tim pada perusahaan Kudo memakai teknologi ini dalam pekerjaannya

Penulis diperintahkan oleh atasan untuk memakai teknologi ini. Menurut penulis, teknologi ini sangat berguna untuk mengontrol revisi pada pekerjaan penulis. Penulis juga memakai teknologi ini untuk berkolaborasi dengan rekan penulis yang mengerjakan pekerjaan yang sama dengan penulis.

2. E-mail dan Slack

Dalam berkomunikasi pada saat kerja, penulis mengandalkan E-mail dan Slack. E-mail pada perusahaan Kudo menurut penulis biasanya digunakan untuk pengingat atau pengumuman suatu *meeting* dan pekerjaan. Slack biasa penulis gunakan untuk komunikasi tentang pekerjaan atau hal non-formal lainnya kepada karyawan atau mahasiswa yang magang pada perusahaan Kudo.

3. Python versi 2.7

Penulis menggunakan bahasa Python untuk mengerjakan pekerjaannya. Python adalah bahasa yang *interpreted, object-oriented, high-level programming language* dengan semantik yang dinamis [15]. Dalam mengerjakan pekerjaan penulis yang berhubungan dengan data, penulis memakai beberapa *package*. *package* yang penulis gunakan untuk mengolah data antara lain: *sk-learn*, *pandas*, *numpy*, dan *scipy*. Penulis menggunakan *package flask* untuk membuat RESTful API.

4. Jupyter Notebook

Jupyter Notebook adalah aplikasi website yang digunakan untuk membuat dokumen yang berisi *live code*, *equations*, visualisasi, dan penjelasan teks [16]. Penulis

menggunakan teknologi ini untuk hasil dan visualisasi data yang sudah penulis *code* dan membagikannya ke rekan dan mentor penulis untuk dianalisis.

5. Sublime Text

Sublime Text adalah suatu *text editor* yang dilengkapi dengan beberapa fitur yang canggih. Sublime Text digunakan untuk *code*, *markup*, dan *prose* [17]. Penulis menggunakan teknologi ini untuk menulis modul *recommender system* yang penulis buat dengan bahasa Python. Penulis merasa bahwa *text editor* ini mampu meningkatkan produktivitas penulis dalam melakukan *coding*.

6. Pycharm IDE

Pycharm adalah *Integrated Development Environment* yang digunakan untuk *programming* bahasa Python. Pycharm dibuat oleh perusahaan JetBrains [18]. IDE ini mempunyai fitur untuk analisis kode, *debug*, *unit test*, Integrasi dengan *version control*, dan *web development*. Penulis menggunakan teknologi ini untuk merapihkan kualitas dari kode yang dibuat. Penulis juga menggunakan IDE ini untuk melakukan *debug* pada program yang penulis buat.

7. Laravel

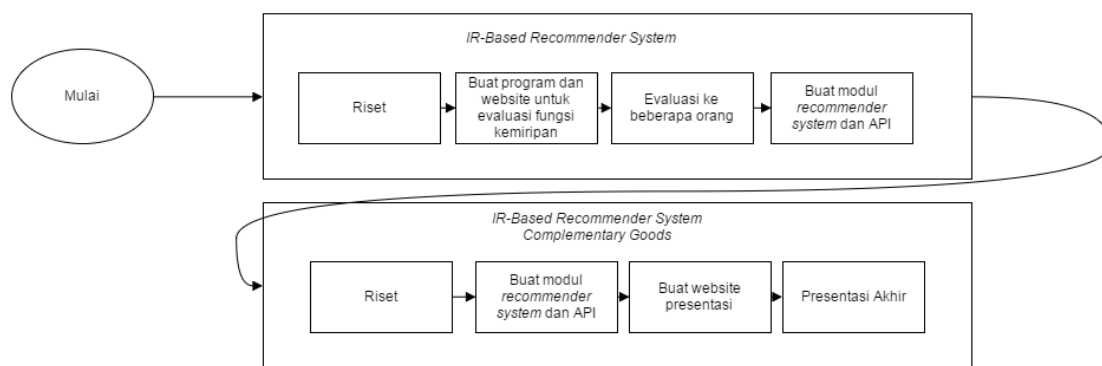
Laravel adalah *web framework* bahasa PHP yang gratis dan *open source* yang memakai model MVC *pattern* [19]. Penulis menggunakan *framework* ini untuk membuat website yang melakukan visualisasi terhadap *recommender system* yang penulis buat. Website ini akan ditujukan pada presentasi penulis.

2.1.5. Gambaran Pekerjaan

Selama kerja praktik, penulis fokus pada mengerjakan rekomendasi produk dengan teknik *content-based recommender system* yang didalamnya merupakan teknik *Information Retrieval*. Adapun fitur yang penulis buat pada kerja praktik dengan teknik ini dibagi menjadi dua, yaitu *recommender system* untuk pencarian

rekomendasi barang yang mirip dan rekomendasi barang yang komplemen terhadap barang yang dilihat oleh pengguna. Fitur yang penulis kerjakan terlebih dahulu yaitu *recommender system* untuk pencarian rekomendasi barang yang mirip terhadap barang yang dilihat oleh pengguna.

Pada bagian ini akan dijelaskan lebih detail pengaplikasian metodologi yang sudah dijelaskan sebelumnya. Gambar 3 memperlihatkan alur kerja penulis melaksanakan kerja praktik di Kudo yang sudah dipaparkan di bagian Metodologi.



Gambar 3 Alur gambaran kerja penulis

Untuk fitur yang pertama, karena penulis tidak tahu bagaimana cara membuatnya, penulis dan rekan penulis diberikan waktu untuk melakukan penelitian terhadap teknik ini. Penulis mendapat beberapa tulisan mengenai *content-based recommender system*. Penulis menemukan bahwa pekerjaan penulis bukan *content-based recommender system* namun merupakan masalah *information retrieval* yang bisa direpresentasikan dalam bentuk VSM. Salah satu pembobotan populer pada VSM yang penulis temukan yaitu dengan menggunakan tf-idf. Selain itu, penulis menemukan bahwa untuk melakukan ranking terhadap rekomendasi produk, diperlukan suatu fungsi kemiripan antara satu barang dengan barang yang lain. Disini

penulis menemukan empat fungsi kemiripan, yaitu *cosine similarity*, *jaccard similarity*, *euclidean distance*, dan *chebychev distance*.

Setelah penulis melakukan penelitian, penulis dan rekan penulis diminta untuk mencari fungsi kemiripan mana yang terbaik dari beberapa fungsi kemiripan yang penulis temukan. Untuk mengetesnya, pertama perlu dibuat programnya *recommender system* dengan fungsi kemiripan tersebut. Untuk meringankan pekerjaan, penulis hanya membuat program untuk *cosine similarity* dan *jaccard similarity* saja sedangkan rekan penulis membuat program sisanya. Penulis dan rekan penulis membuat programnya dengan bahasa Python dan menggunakan beberapa *package* dari bahasa python. Untuk *cosine similarity* sudah diimplementasikan pada *package sk-learn*, sehingga penulis hanya perlu memakainya saja. Sedangkan untuk *jaccard similarity*, penulis perlu membuatnya sendiri. Penulis membuat fungsi kemiripan *jaccard similarity* dengan bantuan *package numpy* dan *scipy*. Penulis mengerjakan keduanya dengan bantuan teknologi Jupyter Notebook dan Sublime Text.

Dalam membuat fungsi kemiripan tersebut, penulis menggunakan dataset yang diberikan oleh mentor penulis. Dataset tersebut berisi tentang barang-barang yang tersedia di basis data perusahaan Kudo. Beberapa fitur yang penulis dan rekan penulis gunakan pada dataset tersebut yaitu nama barang, deskripsi barang, kategori, dan harga. Kategori dan harga akan menjadi fitur untuk menyortir barang rekomendasi dimana barang yang kategori yang sama dan harga yang berdekatan akan direkomendasi lebih dahulu dibanding yang lain.

Setelah itu, kami diperintahkan untuk evaluasi fungsi kemiripan tersebut dengan *testing* kepada manusia. Untuk mengetesnya, perlu dibuat suatu wadah untuk bisa memperlihatkan fungsi kemiripan tersebut terhadap *tester*. Kami mendapat ide untuk

membuat suatu website yang dapat melakukan visualisasi terhadap hasil dari *recommender system*nya. Kami menggunakan *flask* dan *laravel* untuk membuat websitenya.

Setelah kami membuat websitenya, kami mengetesnya ke semua orang pada tim data dan beberapa orang pada tim lain. Cara evaluasinya yaitu dengan memilih barang rekomendasi yang relevan terhadap barang yang dilihat dengan menggunakan sebuah fungsi kedekatan. Barang yang relevan akan membuat *score* dari fungsi kedekatan tersebut naik. Jika ada barang yang tidak relevan, maka *score* fungsi kedekatan tersebut akan dikurangi. *Score* yang terdapat pada keempat fungsi kedekatan tersebut akan diurutkan dan dicari fungsi kedekatan yang mempunyai nilai paling tinggi.

Kami mendapat hasil bahwa *cosine similarity* merupakan fungsi kemiripan yang paling baik dibanding yang lain. Setelah kami mendapatkan hasilnya, kami diperintahkan untuk membuat modul untuk program *recommender system* dengan memakai fungsi kemiripan *cosine similarity*.

Modul yang penulis buat memuat fungsi untuk pembuatan model VSM dan fungsi untuk prediksi hasil rekomendasinya. Dalam proses pembuatannya, penulis memanfaatkan hasil *code* yang digunakan pada saat melakukan evaluasi fungsi kemiripan. Modul ini akan dimanfaatkan oleh RESTful API yang dibuat oleh penulis.

Dalam pembuatan RESTful API, penulis juga memanfaatkan RESTful API yang dibuat pada saat evaluasi fungsi kemiripan. Transaksi yang dilayani pada API ini adalah GET untuk mengambil rekomendasi produk dan POST dalam pembuatan model. Transaksi yang dipakai untuk website demo nantinya hanyalah GET untuk mengambil rekomendasi produk.

Setelah penulis dan rekan penulis selesai mengerjakan fitur yang pertama, penulis melanjutkan untuk mengerjakan fitur rekomendasi barang yang komplemen terhadap barang yang dilihat oleh pengguna. Pada saat mengerjakan ini, rekan penulis sudah tidak bersama penulis lagi dalam mengerjakan fitur ini dikarenakan masa kerja praktik rekan penulis sudah berakhir. Oleh karena itu, penulis hanya bekerja sendiri dalam menyelesaikan fitur ini.

Karena penulis kurang tahu cara membuat fitur ini, penulis diperintahkan oleh mentor penulis untuk riset caranya dengan mencari *paper* yang berhubungan. Penulis menemukan *paper* yang berjudul “*Inferring Networks of Substitutable and Complementary Products*” [20] yang menurut penulis relevan dengan pekerjaan penulis. Sayangnya, data yang tersedia di Kudo tidak bisa menggunakan algoritma pada *paper* tersebut.

Penulis tidak menemukan *paper* lain yang relevan dengan fitur ini. Akhirnya, penulis mempunyai ide untuk menggunakan *content-based filtering* dengan melihat kategori barang yang dilihat terhadap kategori barang komplemennya. Sebagai contoh barang laptop akan dipasangkan terhadap kategori aksesoris laptop dan charger laptop. Untuk memasangkan kategori komplemennya, penulis memasangkan kategorinya secara manual.

Setelah penulis memasangkan semua kategori komplemennya secara manual, penulis mulai membuat modul untuk fitur ini. Contoh pemasangan kategori secara manual yang dilakukan penulis digambarkan pada gambar Y. Karena implementasi fitur ini sangat mirip dengan fitur sebelumnya, penulis hanya perlu memodifikasi *code* yang dibuat sebelumnya agar bisa digunakan untuk fitur ini. Yang berbeda dari fitur sebelumnya adalah pada fitur ini hanya akan mencari rekomendasi berdasarkan kategori yang dipasangkan saja.

main_item	related_category		
679	545, 682, 636		
542	545, 682, 636		
545	679, 542, 629		
682	679, 542, 629		
636	679, 542, 629		
584	580, 581, 593		
593	580, 581, 584		
550	1505, 1502		
645	1505, 1502		

Gambar 4 Pemasangan kategori komplemen

Setelah itu, penulis membuat RESTful API untuk fitur ini. API pada fitur ini juga mirip dengan fitur sebelumnya, dimana transaksi yang dilayani pada API ini adalah GET untuk mengambil rekomendasi produk dan POST dalam pembuatan model. Untuk mengerjakannya, penulis mengubah sedikit *code* pada API yang digunakan untuk fitur sebelumnya.

2.2. Analisis

2.2.1. Pelaksanaan Kerja Praktik

Dalam pelaksanaan kerja praktik di perusahaan Kudo, terdapat perbedaan antara Kerangka Acuan Kerja Praktik (KAKP) dengan pelaksanaan kerja praktik. Terdapat perbedaan tanggal yang terdapat pada KAKP dengan log kerja praktik disebabkan penulis sakit beberapa hari. Selain itu, penulis hanya menuliskan jika penulis akan mengaplikasikan algoritma *Machine Learning* saja pada fitur *recommender system*. Penulis tidak menuliskan bahwa masalah ini diselesaikan dengan mengaplikasikan teknik *Information Retrieval*.

Perbedaan lain pada KAKP penulis yaitu tentang judul KP yang memuat implementasi algoritma *content-based filtering*. Penulis baru mengetahui setelah

penulis melakukan *research* tentang hal ini. Pekerjaan yang diberikan kepada penulis dikerjakan dengan menggunakan teknik *Information Retrieval*. Penulis tidak memakai teknik *content-based filtering* pada pekerjaan yang penulis kerjakan pada saat melaksanakan kerja praktik.

Kendala utama yang penulis alami pada saat pelaksanaan kerja praktik yaitu tidak ada yang mengajarkan penulis tentang fitur *recommender system*. Hal ini disebabkan karena tim *Data Engineer* di perusahaan Kudo baru saja terbentuk dan mengambil pekerja dari bidang lain, sehingga tidak ada yang sangat ahli di bidang *Data Engineer*. Untuk mengatasinya, penulis harus belajar sendiri tentang pengembangan fitur ini.

Karena tidak ada yang ahli di bidang ini, tidak ada yang tahu pasti bagaimana cara mengevaluasi *recommender system* dengan baik. Evaluasi yang penulis lakukan merupakan evaluasi yang dibuat dari ide mentor penulis, penulis, dan rekan penulis. Tentunya hasil evaluasi ini belum tentu hasil yang terbaik dikarenakan belum pernah diuji oleh orang yang ahli.

Selain itu, kendala yang penulis rasakan yaitu keterbatasan akses basis data Kudo. Penulis tidak diberikan akses ke basis data Kudo. Hal ini menyebabkan keterbatasan kreativitas penulis dalam mengolah data untuk fitur yang penulis kerjakan. Untuk mendapatkan datanya, penulis harus meminta data kepada mentor penulis.

Penulis merasa bahwa hari kerja penulis lebih banyak dipakai untuk membuat website demo dibandingkan dengan pekerjaan yang berkaitan dengan data. Penulis merasa bahwa pekerjaan ini bukan merupakan *jobdesk* dari penulis. Namun karena tidak ada yang bisa diminta tolong untuk website demo, penulis terpaksa untuk membuat website demo dimana pembuatan websitenya memakan waktu yang lama.

Terlepas dari kendala dan keluhan yang penulis rasakan, penulis merasa nyaman dalam melaksanakan kerja praktik di perusahaan Kudo. Lingkungan kerja yang terbuka dan bebas membuat penulis betah untuk bekerja di perusahaan ini. Penulis bisa datang ke kantor secara bebas. Perusahaan Kudo juga merupakan tempat yang baik untuk belajar. Karyawan perusahaan Kudo sangat terbuka untuk berbagi ilmunya. Selain itu, ada seminar yang diadakan tiap minggu yang penulis bisa ikuti. Mentor yang sangat responsif juga salah satu hal yang membuat penulis merasa nyaman melaksanakan kerja praktik disini.

2.2.2. Relevansi dengan Perkuliahan di Fasilkom UI

Terdapat beberapa relevansi mata kuliah di Fasilkom UI dengan pekerjaan kerja praktik. Mata kuliah yang relevan dengan kerja praktik yang penulis lakukan di Kudo antara lain Information Retrieval (IR), Perancangan dan Pemrograman Web (PPW), dan Dasar-Dasar Pemrograman (DDP).

Pada mata kuliah DDP, penulis mempelajari bagaimana cara pembuatan *class* pada modul yang penulis buat. Meskipun bahasa yang diajarkan pada DDP menggunakan bahasa Java, tetapi konsep pembuatan *class* dengan memakai prinsip *Object Oriented Programming* (OOP) di Java mirip dengan Python. Prinsip *loop*, *decision*, *array* yang diajarkan di DDP juga membantu penulis dalam melaksanakan pekerjaan praktik.

Selanjutnya mata kuliah PPW membantu penulis untuk membuat website demo. *Web Framework*, *Cascading Style Sheet* (CSS), Javascript, dan *Hypertext Markup Language* (HTML) sangat membantu penulis dalam melaksanakan kerja praktik. Dengan pengetahuan yang penulis dapatkan setelah mengambil mata kuliah PPW, penulis tidak merasa kesulitan dalam membuat website demo.

Terakhir, mata kuliah IR adalah kunci utama dari pekerjaan yang penulis kerjakan. Pada saat pelaksanaan kerja praktik, penulis belum mengambil mata kuliah ini.

Materi tentang evaluasi dan *Vector Space Model* sangat dipakai dalam pekerjaan penulis. Jika penulis sudah mengambil mata kuliah ini, penulis dapat mengatasi kendala tentang evaluasi *recommender system* yang penulis buat.

2.2.3. Pembelajaran *Soft Skill* dari Kerja Praktik

Beberapa kemampuan *soft skill* yang penulis dapatkan pada kerja praktik di perusahaan Kudo antara lain, kemampuan presentasi, komunikasi, dan *problem solving*.

Kemampuan presentasi merupakan kemampuan yang penulis anggap penting pada kerja praktik di perusahaan Kudo. Penulis dituntut untuk dapat menjelaskan fitur yang penulis kerjakan dengan cara presentasi didepan karyawan Kudo. Pada saat penulis melakukan presentasi didepan petinggi Kudo, Penulis merasa bahwa penulis melakukan presentasi dengan kurang baik. Penulis mendapat beberapa nasihat dari beberapa orang yang menonton presentasi penulis. Disini, penulis mendapat banyak hal baru tentang bagaimana cara presentasi yang baik.

Kemampuan komunikasi juga merupakan hal nonteknis penulis dapatkan. Pada saat hari kerja, beberapa anggota tim data sering menanyakan kepada penulis tentang *progress* pekerjaan pada saat pertemuan tim data. Penulis harus menjelaskan tentang pekerjaan penulis ke penanya penulis dengan baik. Selain itu, untuk menjelaskan fitur yang penulis buat kepada tim yang lain, penulis harus bisa menerjemahkan bahasa teknis ke bahasa yang dapat dimengerti oleh orang non-IT.

Tidak jarang penulis dan rekan penulis dipanggil oleh mentor untuk berkumpul membahas tentang masalah pekerjaan. Penulis perlu memikirkan jalan keluar yang sesuai untuk memecahkan masalah ini. Disinilah penulis perlu kemampuan untuk *problem solving*. Banyaknya masalah pekerjaan yang timbul pada saat kerja praktik di perusahaan Kudo menaikkan kemampuan untuk *problem solving* dari penulis.

BAB 3

PENUTUP

3.1. Kesimpulan

Penulis telah melaksanakan kerja praktik selama sepuluh minggu mulai 13 Juni 2016 sampai 19 Agustus 2016. Penulis telah melaksanakan kerja praktik di salah satu perusahaan *startup* yang bernama Kudo. Sebelum diterima di perusahaan ini, penulis bekerja keras untuk mendapatkan perusahaan yang dapat menerima penulis dalam melaksanakan kerja praktik. Selama bekerja di perusahaan Kudo, penulis ditempatkan pada posisi *Data Engineer* yang berada di bawah divisi *data scientist*.

Penulis mendapat pengalaman kerja mengimplementasikan fitur *recommender system* pada aplikasi perusahaan Kudo. Penulis mendapatkan pengalaman berharga kerja praktik di perusahaan Kudo yang memakai teknologi yang terkait dengan data. kemampuan komunikasi, presentasi, dan *problem solving* penulis juga terlatih dalam pelaksanaan kerja praktik.

Terlepas dari kendala dan keluhan yang penulis dapatkan, penulis merasa nyaman dalam pelaksanaan kerja praktik di perusahaan Kudo. Penulis mempunyai kebebasan dalam memperoleh ilmu disini. Penulis merasa bersyukur dapat melaksanakan suasana lingkungan perusahaan *startup* yang cenderung santai namun bergerak cepat. Pekerjaan yang ditawarkan juga terdapat relevansi dengan mata kuliah yang diajarkan di perkuliahan di Fasilkom UI.

3.2. Saran

Melaksanakan kerja praktik akan mendapatkan pengalaman banyak terutama mendapatkan pengetahuan hal yang baru. namun untuk mendapatkannya, penulis menyarankan kepada calon pengambil mata kuliah kerja praktik untuk menyiapkan

diri untuk mencari tempat kerja praktik. Perlu diperhatikan tata cara mendaftar magang pada suatu perusahaan. Jangan mengulur waktu dalam melamar kerja praktik di suatu perusahaan. Calon pengambil mata kuliah kerja praktik harus gencar mencari informasi lowongan kerja praktik. Sebelum dipanggil *interview*, penulis menyarankan untuk menyiapkan diri sehingga pada saat *interview* tidak terjadi kesalahan bicara. Persiapan diri merupakan hal yang penting untuk dapat diterima magang di perusahaan.

Penulis juga menyarankan kepada calon pengambil mata kuliah kerja praktik untuk mendaftar bidang yang sesuai dengan minat. Penulis berharap dengan mendaftar bidang yang sesuai minat, calon pengambil mata kuliah kerja praktik dapat melakukan eksplorasi mengenai pekerjaan bidang tersebut. Calon pengambil mata kuliah kerja praktik dapat mengetahui gambaran tentang lingkungan kerja sebenarnya dari bidang yang ia senangi. Kerja praktik dilakukan agar calon pengambil mata kuliah kerja praktik tidak terkejut dengan dunia kerja sebenarnya jika mengeluti bidang yang disenangi..

DAFTAR REFERENSI

- [1] "Profil Perusahaan - Kudo." *Kudo*. N.p., n.d. Web. 22 Nov. 2016.
- [2] M. Khari, A. Jain, S. Vij and M. Kumar, "Analysis of various information retrieval models," 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2016, pp. 2176-2181.
- [3] Ricci, Francesco, Lior Rokach, and Bracha Shapira. "Introduction to Recommender Systems Handbook." *Recommender Systems Handbook* (2010): 1-35. Web.
- [4] Melville, Prem, and Vikas Sindhwani. "Recommender systems." Encyclopedia of machine learning. Springer US, 2011. 829-838.
- [5] Turney, Peter D., and Patrick Pantel. "From frequency to meaning: Vector space models of semantics." *Journal of artificial intelligence research* 37.1 (2010): 141-188.
- [6] Salton, Gerard, Anita Wong, and Chung-Shu Yang. "A vector space model for automatic indexing." *Communications of the ACM* 18.11 (1975): 613-620.
- [7] Fautsch, Claire, and Jacques Savoy. "Adapting the tf idf vector-space model to domain specific information retrieval." *Proceedings of the 2010 ACM Symposium on Applied Computing*. ACM, 2010.
- [8] Qian, Gang, et al. "Similarity between Euclidean and cosine angle distance for nearest neighbor queries." *Proceedings of the 2004 ACM symposium on Applied computing*. ACM, 2004.
- [9] Niwattanakul, Suphakit, et al. "Using of Jaccard coefficient for keywords similarity." *Proceedings of the International MultiConference of Engineers and Computer Scientists*. Vol. 1. 2013.

- [10] Barrett, Paul. "Euclidean distance: Raw, normalised, and double-scaled coefficients." Unpublished paper retrieved from http://www.pbmetrix.com/techpapers/Euclidean_Distance.pdf (2006).
- [11] Abello, James, Panos M. Pardalos, and Mauricio GC Resende, eds. Handbook of massive data sets. Vol. 4. Springer, 2013.
- [12] Rouse, Margaret. "What Is RESTful API? - Definition from WhatIs.com." *SearchCloudStorage*. N.p., n.d. Web. 22 Nov. 2016.
- [13] "What Is Complementary Good? Definition and Meaning." *BusinessDictionary.com*. N.p., n.d. Web. 22 Nov. 2016.
- [14] Atlassian. "Bitbucket | The Git Solution for Professional Teams." *Bitbucket*. N.p., n.d. Web. 22 Nov. 2016.
- [15] "Welcome to Python.org." *Python.org*. N.p., n.d. Web. 22 Nov. 2016.
- [16] "Project Jupyter." *Project Jupyter*. N.p., n.d. Web. 22 Nov. 2016.
- [17] "Sublime Text." *Sublime Text: The Text Editor You'll Fall in Love with*. N.p., n.d. Web. 22 Nov. 2016.
- [18] Taft, Darryl K. "JetBrains Strikes Python Developers with PyCharm 1.0 IDE." *Technology News, Tech Product Reviews, Research and Enterprise Analysis - EWeek.com*. N.p., n.d. Web. 22 Nov. 2016.
- [19] Gafitescu, Daniel. "Goodbye CodeIgniter, Hello Laravel." *SitePoint*. N.p., 2013. Web. 22 Nov. 2016.

- [20] McAuley, Julian, Rahul Pandey, and Jure Leskovec. "Inferring networks of substitutable and complementary products." Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2015.
- [21] Greengrass, Ed. "Information retrieval: A survey." (2000).

LAMPIRAN 1
KERANGKA ACUAN KERJA PRAKTIK

LAMPIRAN 2
LOG KERJA PRAKTIK