

QUESTION/ANSWERING DENGAN ANSWER QUALITY PREDICTOR NON-TEXTUAL FEATURE

Fajar Baskoro¹, Dwi Sunaryono²
^{1,2}Jurusan Teknik Informatika
 Fakultas Teknologi Informasi
 Institut Teknologi Sepuluh Nopember Surabaya
 Email: fajarbaskoro@gmail.com¹, dwi@its-sby.edu²

Abstrak. *Question/Answering adalah sebuah aplikasi berbasis web dimana setiap user dapat mengajukan pertanyaan dan menjawab pertanyaan dan berpartisipasi didalam mengevaluasi pertanyaan dan / atau jawaban dalam berbagai topik. Kebebasan dan kemudahan dalam partisipasi user didalam sistem ini, menyebabkan beragamnya kualitas dari informasi yang ada didalam sistem. Di dalam paper ini dikembangkan sebuah aplikasi Question/Answering Portals yang dimana user dapat mengajukan pertanyaan, menjawab pertanyaan, melihat jawaban yang diberikan oleh user lain, mengevaluasi jawaban user lain, dan melakukan pencarian terhadap pertanyaan / jawaban yang sudah terselesaikan. Aplikasi ini juga dilengkapi dengan answer quality predictor yang menerapkan metode non-textual feature. Berbeda dengan aplikasi-aplikasi yang sebelumnya telah ada, aplikasi yang dikembangkan ini memiliki kemampuan untuk melakukan penilaian terhadap jawaban yang diberikan oleh user. Dalam melakukan penilaian terhadap jawaban, answer quality predictor menggunakan fitur-fitur non-textual yang pada sebuah jawaban seperti panjang jawaban, rekomendasi user, jumlah klik, dan beberapa fitur lainnya.*

Kata Kunci: *social media, question/answering portal, quality predictor, non-textual feature*

Pada awal perkembangan internet, sebagian informasi yang ada pada internet dibuat dan disusun oleh *publisher*. Informasi ini dikenal dengan sebutan *traditional content*. Mayoritas pengguna internet hanya berperan sebagai *consumer* informasi tersebut. Namun, sejak awal tahun 2000, *user-generate content* menjadi sangat populer dalam dunia internet, semakin banyak pengguna internet yang berpartisipasi dalam penyediaan informasi dan bukan hanya sebagai *consumer* semata. *User-generated content* atau disebut juga *social media* yang populer antara lain adalah *blog*, *web forum*, serta *photo and video sharing community*.

Question/Answering System (Sistem Tanya/Jawab) merupakan salah satu *user-generated content* yang banyak diakses oleh pengguna internet saat ini. Pada sistem ini, *user* berperan untuk menjawab pertanyaan yang diajukan oleh *user* lain. Sistem ini menjadi alternatif untuk mendapatkan informasi pada website. Dibandingkan dengan mencari melalui *search engine*, dimana hasilnya tidak selalu sesuai dengan apa yang dibutuhkan oleh *user*, dalam *Question/Answering System*, *user* memberikan informasi yang dibutuhkan oleh *user* lain yang bertanya.

User-generate content / Social media bisa digunakan sebagai media untuk

mendapatkan informasi, tetapi ada perbedaan yang sangat penting antara *traditional content* dan *user-generate content*. Perbedaan itu terletak pada kualitas informasi yang diberikan antara keduanya. Pada *traditional content* biasanya kualitas informasi yang diberikan selalu baik dan dapat dipercaya. Sedangkan pada *social media* kualitas informasi yang diberikan sangat beragam, mulai dari *high-quality*, *low-quality* bahkan kita dapat menjumpai informasi yang hanya berupa *spam*. Hal ini karena pada *social media* website, berbagai macam user ikut serta memberikan informasi. Oleh karena itu dalam paper ini akan dibahas suatu *Question/Answering System* yang didalamnya terdapat proses pengidentifikasian kualitas informasi dari jawaban – jawaban yang diberikan oleh *user* dengan menggunakan metode *non-textual feature*.

Social Media

Social Media adalah aplikasi berbasis internet untuk sharing dan berdiskusi tentang informasi dengan sesama. Social Media sering dikaitkan dengan teknologi yang terintegrasi, interaksi sosial, serta pembentukan dan pertukaran kata, gambar, video dan audio [5]. Pada social media informasi yang diberikan bergantung pada banyak perspektif dan

komunitas sebagai tempat untuk melakukan sharing cerita dan pengalaman.

Social media sangat berbeda dari industrial media seperti koran, televisi, dan film. *Social media* adalah alat yang relatif murah bagi semua orang untuk menerbitkan atau mengambil informasi.

Social media memiliki banyak bentuk yang berbeda antara lain: Internet Forum, Web Blog, Social Networking, Social Bookmarking, Social New Site, Wikis, Image Hosting Service, Audio and Video Hosting Service, dan Question/Answering Portal.

Salah satu contoh Question/Answering yang saat ini populer adalah Yahoo! Answer.

Question / Answering Portals

Question/Answering portals and forums adalah sistem dimana tiap *user* dapat mengajukan pertanyaan dan menjawab pertanyaan dalam berbagai topik[2]. Hal yang membuat sistem ini menarik adalah *user* dapat berinteraksi dan mendapatkan informasi dari berbagai macam user. Disini semua user akan membentuk suatu jaringan sosial dengan interaksi heterogen didalamnya.

Question/Answering Portals (Q/A) berbeda dengan *Question Answering System (QAS)*, dalam QAS sistem melakukan *question answering* secara otomatis, sedangkan dalam sistem Q/A merupakan suatu service dimana pengguna yang akan menjawab pertanyaan dari pengguna lainnya.

Dalam *Question/Answering portals* user tidak hanya melakukan tanya – jawab, tetapi juga ikut mengevaluasi. User bisa mem – vote jawaban dari user lain, menandai pertanyaan yang menarik, serta melaporkan penyalahgunaan forum. Secara keseluruhan setiap user memiliki tiga fungsi yaitu: *asker*, *answerer*, dan *evaluator*. Sebagai contoh, jenis – jenis aktivitas user pada Yahoo! Answer. User dapat mengajukan pertanyaan dan memberikan bintang pada pertanyaan yang dianggap menarik. User juga dapat memberikan jawaban, memilih jawaban terbaik pada suatu pertanyaan, serta mengevaluasi jawaban – jawaban yang ada dengan cara merekomendasikan atau mendisrekomendasikan jawaban.

Document Quality Content

Dalam *user-generated-content* atau *social media* dimana pengguna ikut

berpartisipasi dalam memberikan informasi dan pengetahuan, tidak menjamin bahwa kualitas dari informasi yang diberikan oleh pengguna sistem tersebut memiliki kualitas yang baik. Beberapa bahkan memiliki kualitas yang rendah dan dapat dikategorikan sebagai *spam*. Hal ini karena semua orang dari berbagai kalangan dan profesi dapat ikut berpartisipasi memberikan jawaban. Tidak ada jaminan bahwa pengguna yang tidak memiliki keahlian dalam suatu bidang untuk tidak ikut berpartisipasi menjawab pertanyaan pada bidang yang tidak dikuasainya.

Di bawah ini merupakan salah satu contoh bad answer dan goog answer yang dapat ditemui dalam Question/Answerin Portal

Pertanyaan :

Apa yang dimaksud Token Ring?

Bad Answer :

Numpang lewat aja.

Good Answer :

Token ring adalah sebuah skema dalam suatu jaringan komputer lokal (local area network) terutama bagaimana suatu pesan dikirimkan dari suatu computer ke computer lain.

Di dalam jaringan, komputer yang bisa mengirimkan data adalah yang mempunyai tanda (token). Token ini kemudian diberikan kepada komputer tetangganya yang terdekat dan terus berputar. Token ini akan terus berputar, walaupun tidak ada berita yang dikirimkan.

Dalam menentukan kualitas informasi (khususnya dalam *Question/Answering Portals*) ada dua metode yang dapat dilakukan yaitu :

User Recommendation

Salah satu fungsi pengguna dalam *Question/Answering Portals* adalah sebagai evaluator. Pengguna dapat melakukan vote untuk jawaban yang dianggap berkualitas, memilih jawaban terbaik dari suatu pertanyaan serta menandai suatu pertanyaan yang menarik. Dari hasil evaluasi pengguna tersebut, kita bisa mendapatkan rekomendasi tentang jawaban – jawaban yang berkualitas.

Namun sayangnya, fungsi ini sering tidak berfungsi dengan baik karena adanya sifat subyektivitas pengguna dan kurangnya kesadaran dari pengguna untuk menjalankan perannya sebagai evaluator.

System Recommendation

System recommendation adalah penilaian sebuah kualitas konten berdasarkan dari data-data. Hasilnya akan lebih akurat karena tidak terdapat unsur-unsur subyektifitas. Pendekatan yang digunakan adalah metode *Non-Textual Feature*.

Non-Textual Feature memanfaatkan feature diluar konteks penulisan yang dapat digunakan untuk mengestimasi apakah suatu dokumen berkualitas atau tidak.

Non-Textual Feature

Non-Textual Feature dibedakan menjadi dua tipe yaitu *Monotomic Feature* dan *Non-Monotomic Feature*. Monotomic feature adalah feature yang kualitas jawabannya dapat diperkirakan dengan melihat nilai atau frekuensi dari feature tersebut. Dengan kata lain, semakin besar nilai featurenya maka semant baik pula kualitas jawabannya[12]. Feature – feature yang merupakan monotonic adalah :

- *Answerer Acceptance Ratio*
persentase jawaban terbaik dibandingkan dengan seluruh jawaban yang dimiliki oleh penjawab.
- *Answer Evaluation*
nilai bintang yang didapat oleh jawaban.
- *Answerer Activity Level*
Seberapa sering pengguna bertanya dan menjawab dalam sistem
- *Answerer Category Specialty*
Jumlah pertanyaan pada suatu kategori, dan jawaban terbaik pada suatu kategori.
- *User Recommendation*
Banyak merekomendasikan sebuah pertanyaan sebagai pertanyaan yang menarik.
- *User Dis-Recommendation*
Banyak dis-merekomendasikan pertanyaan
- *Click Count*
jumlah akses pada sebuah pertanyaan.
- *Number of Answer*
jumlah jawaban yang diberikan pada sebuah pertanyaan.
- *Save Count*
jumlah pengguna yang menyimpan sebuah pertanyaan kedalam foldernya.

Sedangkan non-monotonic feature adalah feature yang kualitas jawabannya tidak dapat diperkirakan dengan hanya melihat nilai mentah dari feature tersebut, karena nilai feature yang tinggi tidak menjamin bahwa

kualitas jawabannya akan semakin baik pula[12]. Feature – feature yang merupakan non-monotonic adalah :

- Jumlah Kata (*Word*)
semua kata yang ada dalam jawaban.
- Jumlah Kalimat (*Sent*)
semua kalimat dari jawaban.

Berdasarkan uraian diatas, maka dalam menentukan kualitas jawaban menggunakan metode *non-textual feature* kita menggunakan 11 *feature*. Tiap fitur tersebut dihitung nilai frekuensinya. Pada tabel dibawah dapat kita lihat tipe – tipe numerik yang digunakan untuk merepresentasikan nilai dari tiap fitur tersebut.

Tabel 1. Daftar feature dan tipe numeriknya

No.	Nama Feature	Tipe
1.	Jumlah Kata	Integer
2.	Jumlah Kalimat	Integer
3.	Answerer Acceptance Ratio	Persen
4.	Answer Evaluation	1, 2, 3, 4, 5
5.	Answerer Activity Level	Integer
6.	Answerer Category Specialty	Integer
7.	User recommendation	Integer
8.	User Dis-Recommendation	Integer
9.	Click Count	Integer
10.	Number of Answer	Integer
11.	Save Count	Integer

Untuk menghitung kualitas jawaban dengan menggunakan metode non-textual feature ini maka feature – feature non-monotonic harus diubah terlebih dahulu menjadi monotonic. Untuk mengubahnya menjadi monotonic feature digunakan *Kernel Density Estimation (KDE)* [12].

$$kde(x) = \frac{1}{N} \sum_{i=1}^N K(x - x_i) \tag{1}$$

$$\text{Dimana, } K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \tag{2}$$

Tahap selanjutnya adalah menghitung nilai score dari keseluruhan feature.

$$score(x) = \frac{1}{Z(x)} \exp \left[\sum_{i=1}^{11} \lambda_i f_i(x) \right] \tag{3}$$

Dimana x merupakan pasangan pertanyaan dan jawaban yang akan dihitung kualitasnya. $f_i(x,y)$ merupakan nilai dari feature ke-i.

$$f_i(x, y) = \begin{cases} kde(xf_i) & \text{jika, feature-i = non-monotonic} \\ xf_i & \text{jika, feature-i = monotonic} \end{cases} \quad (4)$$

Dimana xf_i merupakan raw value dari feature ke-i, $Z(x)$ adalah faktor normalisasi, dan λ_i adalah nilai parameter penentu dari tiap – tiap feature. Nilai parameter penentu tersebut pada aplikasi ini kita gunakan correlation coefficient dari tiap – tiap feature.

Tabel 2. Correlation coefficient dari tiap feature[12]

No.	Nama Feature	Corr
1.	Jumlah Kata	0.4285
2.	Jumlah Kalimat	0.4285
3.	Answerer Acceptance Ratio	0.1837
4.	Answer Evaluation	0.1675
5.	Answerer Activity Level	0.1982
6.	Answerer Category Specialty	0.2103
7.	User recommendation	0.0351
8.	User Dis-Recommendation	-0.0596
9.	Click Count	-0.0085
10.	Number of Answer	-0.0297
11.	Save Count	0.0469

Correlation coefficient dari tiap feature[12]

I. Metodologi

Aplikasi *Question/Answering Portals* yang ada saat ini hanya memiliki proses evaluasi dengan menggunakan *user recommendation*. Pada paper ini aplikasi *Question/Answering Portals* yang akan dibuat akan memiliki dua proses evaluasi, yaitu *user dan sistem recommendation*.

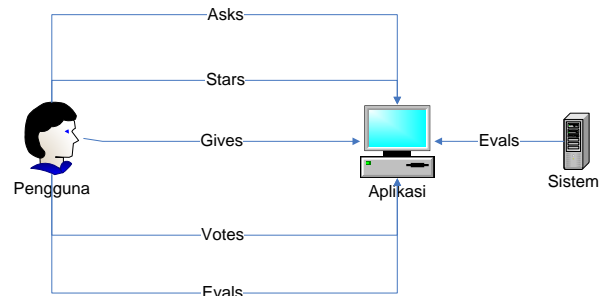
Aplikasi ini merupakan aplikasi berbasis web. Aplikasi ini terdiri dari dua jenis login, yaitu login administrator dan login pengguna. Login administrator adalah login untuk memantau aktivitas pengguna, administrator berhak menghapus pertanyaan atau jawaban yang dianggap melakukan penyalahgunaan pada aplikasi. Sedangkan login pengguna adalah login bagi siapa saja yang ingin menggunakan aplikasi untuk bertanya dan menjawab.

Dalam aplikasi ini pengguna memiliki tiga fungsi utama yaitu *asker, answerer, dan evaluator*.

- *Asker* yaitu pengguna dapat berperan sebagai pemberi pertanyaan.
- *Answerer* yaitu pengguna berperan sebagai pemberi jawaban untuk berbagai pertanyaan yang ada pada aplikasi.

- *evaluator* yaitu pengguna berperan sebagai pemberi penilaian atau evaluasi baik untuk jawaban maupun pertanyaan. Fungsi *evaluator* terdiri dari *Recomendation* pertanyaan, *Voting* jawaban, dan *Evaluation* jawaban.

Fungsi utama pengguna dapat diilustrasikan melalui gambar di bawah ini.



Gambar 1. Rancangan Umum Sistem

Proses evaluasi merupakan proses mengidentifikasi kualitas dari jawaban. Bentuk kualitas jawaban diwakili oleh jumlah bintang. Bintang terendah adalah 1 bintang yang berarti bahwa jawaban memiliki kualitas yang sangat buruk, dan bintang tertinggi adalah 5 bintang, yang berarti bahwa jawaban memiliki kualitas sangat baik. Dalam proses mengidentifikasi kualitas jawaban, sistem ini akan menggunakan dua pendekatan identifikasi kualitas yaitu:

- *User Recommendation*

Untuk melakukan proses memberi bintang pada jawaban, pengguna tinggal memilih berapa bintang yang ingin diberikan untuk suatu jawaban. Bintang terdiri dari bintang satu sampai bintang lima. Semakin banyak bintang yang didapatkan oleh sebuah jawaban mengidentifikasikan bahwa jawaban tersebut dianggap berkualitas oleh pengguna. Pengguna hanya boleh memberikan evaluasinya satu kali pada suatu jawaban. Hasil evaluasi dari semua pengguna pada jawaban tersebut dirata-rata dan akan ditampilkan.

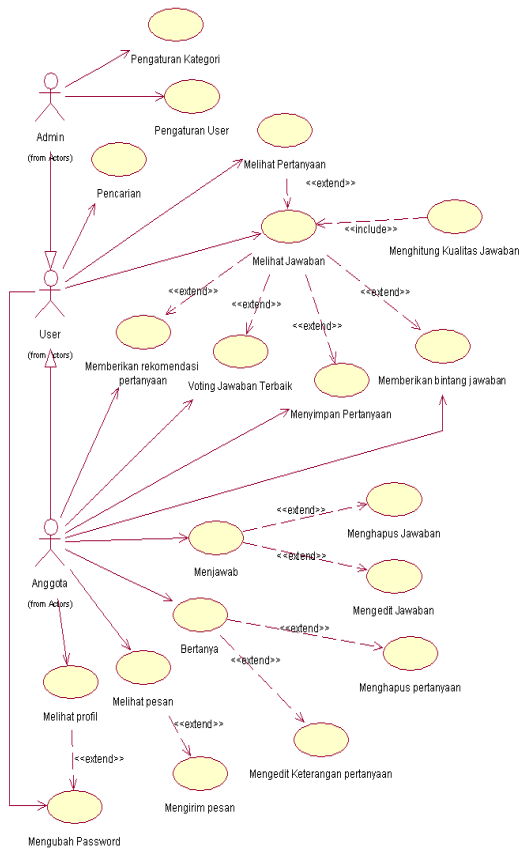
- *System recommendation*

Pada *system recommendation* aplikasi akan menghitung berapa besar kualitas dari sebuah jawaban menggunakan nontextual features yang terdiri dari 11 feature. Feature – feature yang akan digunakan telah dibahas pada bab 2.4. Kualitas jawaban dari sistem ini juga diwakili oleh bintang 1 sampai 5, dimana semakin banyak bintang yang diperoleh maka

semakin bagus kualitas jawaban tersebut menurut sistem.

Perancangan Sistem Use-Case Diagram

Pelaku dalam aplikasi *Question/Answering Portal* ini dibagi menjadi 2 group, yaitu: admin dan anggota. *Use case* pada aplikasi ini terdapat 18 fungsi. *Use case* terbagi menjadi tiga kelompok, yaitu *use case* yang hanya dapat dilakukan oleh aktor admin, *use case* yang hanya dapat dilakukan oleh aktor anggota, dan *use case* yang dapat dilakukan oleh seluruh pengguna pada aplikasi.



Use Case Diagram

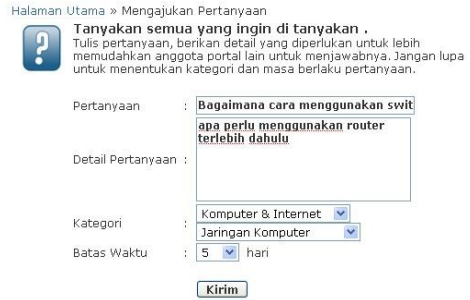
Prediktor digunakan untuk melakukan penilaian terhadap jawaban yang ada didalam sistem. Hasil dari penilaian ini akan ditampilkan dengan bentuk bintang pada bagian penilaian sistem

II. Hasil dan Pembahasan

Hasil uji coba didapatkan dengan menguji jalannya program mulai dari mulai proses *penginputan* data sampai pada *output* data yang diterima.

Uji Coba Fungsi Bertanya

Fungsi bertanya adalah fungsi yang digunakan oleh user untuk mengajukan pertanyaan kedalam aplikasi. Uji coba untuk fungsi bertanya dilakukan dengan cara mengisikan inputan yang ada pada form.



Gambar 2. Uji Coba Bertanya proses Question

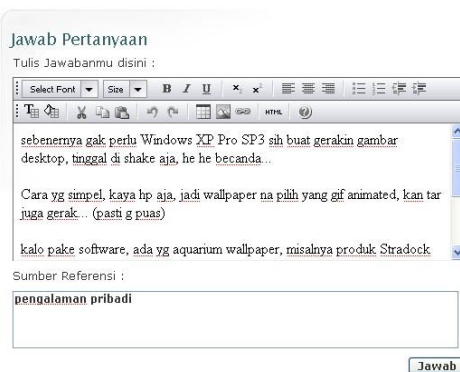
Jika berhasil halaman akan menuju ke halaman pertanyaan atau halaman utama, dimana pertanyaan yang baru saja diinputkan tersebut akan tampil



Gambar 3. Uji Coba Hasil Bertanya

Uji Coba Fungsi Menjawab

Fungsi menjawab merupakan fungsi untuk memberikan jawaban pada pertanyaan yang ada pada aplikasi. Uji coba dilakukan dengan cara memasukkan inputan – inputan pada form.



Gambar 4. Uji Coba Hasil Jawaban



Gambar 5. Hasil Uji Coba Menjawab

Skenario Uji Coba Quality Predictor Sample Data

Pada uji coba *quality predictor* diperlukan sampel data dengan tujuan untuk melakukan pengujian terhadap performa dari *quality predictor* aplikasi. Data yang digunakan untuk uji coba adalah data – data pertanyaan dan jawaban yang ada pada Yahoo! Answer. Data – data yang diambil adalah data – data dari bermacam – macam kategori yang ada mulai dari bisnis, kesehatan, politik, masyarakat, komputer sampai ilmu pengetahuan. Data – data tersebut diambil secara random sebanyak 40 pertanyaan dan 100 jawaban, tanpa membatasi panjang data yang akan diambil. Data – data diambil mulai dari panjang data paling pendek sampai panjang data yang sangat panjang.

Evaluasi Quality Predictor

Evaluasi dilakukan dengan cara membandingkan hasil penilaian user dengan hasil penilaian sistem. Dengan menggunakan perhitungan Precision, Recall dan F-Measure untuk 100 jawaban yang diujicobakan, didapatkan hasil Precision sebesar 45 %, Recall 53 %, dan F-Measure 49 %.

Dari hasil ujicoba *answer quality predictor* tersebut dapat disimpulkan bahwa penilaian kualitas jawaban yang dihasilkan oleh sistem cukup dapat diandalkan dalam memberikan informasi kepada pengguna, jawaban mana yang dapat dipercaya sebagai jawaban yang berkualitas dan dapat dijadikan referensi bagi pengguna untuk menjadi solusi dari pertanyaan yang diajukannya.

Dengan adanya penilaian sistem yang diberikan pada aplikasi *Question/Answering Portal* pengguna dapat dengan mudah memfilter jawaban – jawaban dalam aplikasi dengan bantuan *answer quality predictor* tersebut. Tetapi walaupun penilaian kualitas jawaban yang dihasilkan oleh sistem cukup dapat diandalkan, namun presentasi keakuratannya masih cukup rendah, sebab dari hasil ujicoba dapat diambil kesimpulan bahwa dari 10 jawaban yang ada, terdapat kurang lebih 5 jawaban yang penilaian sistemnya akurat.

III. Simpulan

Setelah dilakukan serangkaian uji coba dan analisa terhadap aplikasi ini, maka dapat disimpulkan sebagai berikut:

- Aplikasi telah berhasil melaksanakan proses bertanya dan menjawab pertanyaan, serta proses – proses lainnya yang mendukung aplikasi ini, seperti proses melakukan voting jawaban terbaik, memberikan bintang pada jawaban, memberikan rekomendasi pertanyaan dan lain-lain.
- Aplikasi telah berhasil melakukan identifikasi kualitas jawaban yang ada pada sistem dengan menggunakan metode non-textual feature. Berdasarkan hasil uji coba yang telah dilakukan, aplikasi telah melakukan ketepatan identifikasi kualitas jawaban dengan nilai precision sebesar 45 % dan recall sebesar 53 % serta F-Measure sebesar 49 %. Dari hasil ujicoba tersebut dapat disimpulkan bahwa penilaian sistem terhadap kualitas jawaban yang ada pada aplikasi cukup dapat diandalkan, walaupun kurang akurat karena dari seluruh jawaban yang ada pada aplikasi hanya terdapat sekitar 50 % jawaban yang diberikan penilaian dengan tepat.

IV. Daftar Pustaka

- [1] Eugene Agichtein , Carlos Castillo , Debora Donato , Aristides Gionis , Gilad Mishne. Finding high-quality content in social media. Proceedings of the international conference on Web search and web data mining, Palo Alto, California, USA, Februari.2013.
- [2] Lada A. Adamic , Jun Zhang , Eytan Bakshy , Mark S. Ackerman. Knowledge sharing and yahoo answers: everyone knows something. Proceeding

- of the 17th international conference on World Wide Web, 2013, Beijing, China, April.2013.
- [3] J. Jeon, W. B. Croft, J. H. Lee. And S. Park. A framework to predict the quality of answer with non-textual features. In SIGIR '06: Proceeding if the 29th annual international ACM SIGIR conference in Reasearch and development in information retrieval, 2014.
- [4] Laxxuss. Advanced Java 2 Platform How To Program, Prentice Hall, 2012.
- [5] <http://www.ilmukomputer.com>. Fundamental of Portal. Ilmu Komputer, Oktober.2008
- [6] Rickyanto,Isak. Belajar Sendiri Java Server Pages. Elex Media Komputindo, 2012.

Halaman ini sengaja dikosongkan.