Implementasi dan Analisis Klasifikasi Spam Pada Pesan Singkat Seluler Dengan Pendekatan Collaborative Filtering Menggunakan Naïve Bayes

Implementation and Analysis of Spam Classification on Short Text Messages Using Collaborative Filtering Approach with Naïve Bayes

Ricky Kristian Butar Butar, Shaufiah, S.T., M.T.[2], Mochamad Arif Bijaksana, Ir, M.Tech[3]

[1,2,3]Prodi S1 Teknik Informatika, Fakultas Teknik, Universitas Telkom

[1]**ricky.kristianb@gmail.com**, [2]**shaufiah@gmail.com**, [3]**arifbijaksana@gmail.com**

Abstrak

Pesan Singkat atau yang dikenal dengan SMS (*Short Message Service*) merupakan layanan pertukaran pesan antar pengguna layanan tersebut. Semakin banyaknya pengguna layanan SMS, tidak sedikit pihak yang memanfaatkannya untuk mendapatkan keuntungan, yaitu dengan menyebarkan SMS sampah, atau dikenal dengan SMS *spam*. Oleh karena itu, pada penelitian tugas akhir ini, penulis melakukan pengklasifikasian terhadap SMS yaitu kelas *spam* maupun *ham*. Pengklasifikasian SMS tersebut dengan menggunakan pendekatan *Collaborative Naïve Bayes* yang berorientasi pada rekomendasi beberapa pengguna dan *Content-Based Naïve Bayes* dengan melihat konten pada SMS. Data rekomendasi didapatkan dengan menyebarkan 300 SMS kepada pengguna. Untuk *Content-Based* dibutuhkan *preprocessing* sehingga konten SMS menjadi seragam, memiliki informasi penting, dan mempercepat proses komputasi. *Preprocessing* yang digunakan adalah *slang handling*, *stopword removal*, dan *stemming*. Pengujian dilakukan dengan membagi SMS menjadi data latih dan data uji sesuai dengan pembagian data *cross validation* yaitu 5-*fold* dan 10-*fold*. Hasil pengujian yang dilakukan menghasilkan tingkat akurasi sebesar 97.12% untuk 5-*fold* dan 97.28% untuk 10-*fold* sehingga dikatakan mampu meningkatkan keakurasian pengklasifikasian SMS dengan menggunakan metode *Collaborative Filtering*.

Kata Kunci : *Pengklasifikasian, Collaborative, Content-Based, Naïve Bayes, Preprocessing*

Abstract

Short messages or more common known as SMS (Short Message Service) is a service for exchanging messages among users. Because of the increasing number of SMS's user, some of the parties exploit to gain advantages for themselves by either spreading or sending junk SMSes, which is known as SMS spam. Hence, this final project is trying to classifyed the SMSes with spam (junk) or ham (legitimate) class. This classification is done by using collaborative naïve-bayes method which is oriented of recommendation from some users and content-based naïve bayes by analyzing the content of a SMS. Recommendation's data are obtained by spreading 300 SMSes to some users. Preprocessing is needed for content-based naïve bayes in order to get uniform content, have useful information, and to expedite the computation. Slang handling, stopword removal, and stemming are used for Preprocessing. SMSes are divided into training set and testing set according to 5-fold and 10-fold data's selection method. From this experiment, the average result for 5-fold is 97.12% and for 10-fold is 97.28% so from the result shows the increasing of SMS classification accuracy using Collaborative Filtering method.

Keywords : *Classification, Collaborative, Content-based, Naïve bayes, Preprocessing.*

1. Pendahuluan

Teknologi komunikasi digunakan untuk dapat saling berkirim interaksi satu dengan yang lain. Telepon selular merupakan contoh teknologi komunikasi yang banyak digunakan oleh masyarakat diseluruh dunia dalam beberapa dekade belakangan ini dan di Indonesia pengguna telepon seluler sekitar 152 juta pengguna pada tahun 2009 [7]. Kepopuleran telepon seluler membawa salah satu layanannya tersebut naik ke level yang lebih tinggi yaitu menjamurnya penggunaan pesan singkat atau yang lebih dikenal dengan SMS (*Short Message Service*). Kelebihan SMS yang ditawarkan yaitu merupakan layanan telepon seluler yang berfokus pada pertukaran pesan tulisan yang murah [4].

Pengguna yang banyak dan kelebihan yang ditawarkan oleh layanan SMS dimanfaatkan oleh beberapa pihak untuk mendapatkan keuntungan baik secara personal maupun setingkat industri, seperti pengiriman pesan iklan komersial, penipuan pengisian pulsa, pemenang undian atau transaksi keuangan palsu kepada pengguna layanan SMS [17]. SMS tersebut dapat dikategorikan sebagai SMS sampah (SMS *Junk* atau *Spam*) yang mengakibatkan terganggunya pengguna telepon seluler seperti selalu munculnya notifikasi SMS *spam* yang dianggap seharusnya tidak diperlukan dan juga berefek terhadap kepadatan trafik operator seluler [6].

Pengklasifikasian SMS merupakan langkah yang dilakukan untuk menyaring SMS menurut kelasnya. *Collaborative Naïve Bayes* dan *Content-based Baïve Bayes* digunakan sebagai metode untuk mengklasifikasikan SMS. *Collaborative Naïve Bayes* merupakan teknik pengklasifikasian berdasarkan opini beberapa *user* untuk

memprediksi kelas SMS dari seorang *user*. [10]. Ketika SMS tidak mempunyai kelas dan belum ada user yang merekomendasikan kelas SMS tersebut, maka *Content-Based Naïve Bayes* digunakan untuk membantu mengklasifikasikan SMS dan menghindari *sparse* yang terlalu besar. SMS tidak memiliki aturan umum, sehingga SMS memiliki beberapa karakteristik seperti SMS yang tidak teratur, tidak ada singkatan, dan SMS formal, oleh karena itu *preprocessing* digunakan untuk membersihkan dan membuat SMS menjadi lebih teratur sehingga informasi dari SMS dapat digunakan untuk mengklasifikasikan SMS. *Preprocessing* yang digunakan antara lain penanganan singkatan (*slang replacing*), merubah kata menjadi kata dasar (*stemming)*, dan menghapus kata yang tidak informatif (*feature extraction* dan *selection).*

Dari sebuah penelitian sebelumnya, didapatkan akurasi dari *Collaborative Naïve Bayes* sebesar 68.5% untuk 50 user [5] dan 98.7% pada *Content-Based Naïve Bayes*[4].

2. Penelitian Terkait

Pada penelitian sebelumnya, pengklasifikasian dengan metode *Collaborative Filtering* digunakan untuk mendapatkan rekomendasi pada *item* (barang) baru berdasarkan opini dari *use*r yang memiliki ketertarikan terhadap *item* yang sama [5][10]. Berbeda dengan penelitian sebelumnya, pada penelitian Tugas Akhir ini, *Collaborative Filtering* digunakan untuk memberikan rekomendasi kelas untuk pesan singkat (SMS) dari *user* yang berbeda-beda dengan melihat probabilitas rekomendasi kelas yang diberikan oleh *user* lainnya terhadap SMS baru yang didapatkan oleh *user* tertentu. Penelitian ini menunjukkan bahwa penggunakan metode *Collaborative Filtering* untuk mengklasifikasikan SMS terbukti dapat digunakan dilihat dari keakurasian yang dihasilkan.

3. Landasan Teori

3.1 *Short Message Service* (SMS)

Pesan singkat atau yang lebih dikenal dengan SMS (*Short Message Service*) merupakan layanan telepon seluler yang memungkinkan pengguna saling bertukar pesan tulisan. SMS umumnya memilki maksimal 160 karakter untuk satu kali pengiriman pesan yang dikirimkan secara *wireless* tanpa membutuhkan jaringan internet di dalamnya [3].

3.2 *Short Message Service Spam*

SMS *spam* dikarakteristikkan dengan dua atribut, pertama adalah pesan yang tidak diinginkan (*unsolicited*), kedua adalah pengiriman dalan jumlah yang sangat banyak (*sent in bulk*)[14]. Karakteristik SMS *spam* sangat berbeda dengan karakteristik pada *email span* dimana jumlah karakter dan konten yang jelas memang terdapat perbedaan dimana pada *email*, *spam* bisa dikategorikan melalui *header*, *subject*, pengirim, dan juga karakter yang dapat ditampung oleh *email* lebih besar daripada SMS yang hanya dapat menampung 160 karakter untuk satu kali pengiriman pesan.

3.3 *Collaborative Filtering*

Collaborative Filtering merupakan metode *filtering* yang paling popular, banyak di implementasikan dan teknik yang paling banyak direkomendasikan [10]. *Collaborative filtering* pada dasarnya bekerja menurut asumsi dari pengguna sistem dengan keinginan yang sama yang dengan memilih keputusan yang sama. Untuk menghasilkan sebuah rekomendasi, *collaborative filtering* pada awalnya membentuk sebuah sistem ketetanggaan dari pengguna dengan tingkat kesamaan yang tinggi antar pengguna. Kemudian menghasilkan sebuah prediksi dengan melakukan kalkulasi bobot penilaian dari pengguna dalam satu ketetanggaan [5].

3.4 *Preprocessing* Data

Pre-processing dilakukan untuk mengubah pesan pada SMS menjadi format standar yang dapat dimengerti oleh *learning machine*. Data yang ingin dilatih kemungkinan besar tidak bersih, seperti adanya *noise*, *missing value*, *outlier* yang tidak sesuai, *inconsistent*, *incomplete*, selain itu, penghapusan elemen-elemen yang diperkirakan tidak berhubungan dilakukan. *Preprocessing* yang dilakukan adalah:

1. *Slang Handling*

   SMS tidak memiliki format atau aturan penulisan didalamnya sehingga pengguna dengan serta merta menulis dengan bebas dan tidak teratur. *Slang Handling* digunakan sebagai penanganan masalah tersebut sehingga SMS menjadi lebih teratur. Penanganan terhadap *slang* tersebut didukung dengan kamus *slang* sebagai acuan perubahan kata-kata pada SMS. Dilakukan pengecekan per kata pada SMS dengan kata pada kamus *slang* dan mengganti kata yang diindikasikan *slang* dengan kata (kata-kata) baru dari kamus *slang*.
2. *Stopword Removal*

   Stopword Removal digunakan untuk menghilangkan kata-kata pada SMS yang dianggap tidak memiliki makna yang biasanya muncul dalam jumlah yang besar. Penanganan *stopword* menggunakan kamus acuan untuk menghilangkan kata.
3. *Stemming*

   Stemming merupakan penganan untuk menemukan kata dasar dari sebuah kata dengan menghilangkan imbuhan (*affixes*) baik yang terdiri dari awalan (*prefixes*), sisipan (*infixes*), dan akhiran (*suffixes*) [8].
4. *Feature Extraction* dan *Selection*

Pesan yang akan dilakukan pengklasifikasian pada umumnya memiliki karakteristik seperti adanya *noise* pada data. Untuk mempelajari suatu data adalah dengan menentukan fitur-fitur yang menjadi informasi berguna dengan cara penyederhanaan isi suatu pesan. *Feature Extraction* dan *Feature Selection* adalah metode untuk memilih informasi kata pada SMS yang dinilai mempuyai informasi yang berguna untuk proses klasifikasi[8]. Pemilihan kata pada SMS dilakukan dengan melihat frekuensi kemunculan kata pada sebuah SMS dan frekuensi kemunculan kata pada kumpulan SMS (dokumen-dokumen).

5. *Term Frequency-Inverse Document Frequency* (TF-IDF)

TF-IDF merupakan dua metode pembobotan yaitu *Term Frequency* dan *Inverse Document Frequency*. TF-IDF biasa digunakan untuk pembobotan pada *information retrieval* dan *text mining*. Nilai dari TF-IDF akan bertambah apabila jumlah kemunculan kata pada dokumen semakin banyak. Nilai dari TF-IDF dapat menentukan kata yang selalu muncul pada kumpulan dokumen dan kata yang jarang muncul pada dokumen [8].

a. *Term Frequency* (TF) : *Term Frequency* digunakan untuk melihat frekuensi kemunculan kata pada sebuah SMS
b. *Inverse Document Frequency* (IDF) : *Inverse Document Frequency* merupakan pembobotan sebuah kata untuk mengukur tingkat kepentingan dari sebuah kata pada kumpulan dokumen (SMS).

6. *Naïve Bayes Classifier*

Naïve Bayes Classifier (NBC) merupakan sebuah pengklasifikasi probabilitas sederhana yang mengaplikasikan Teorema Bayes dengan asumsi ketidaktergantungan (independent) yang tinggi [17].

a. *Collaborative Naïve Bayes*: Pengklasifikasian dengan menggunakan *Collaborative Naïve Bayes* yaitu dengan menghitung probabilitas dari setiap kelas yang di rekomendasikan oleh setiap *user*.
b. *Content-Based Naïve Bayes*: Melihat beberapa rumus diatas maka untuk menentukan kelas SMS pada *Content-Based Naïve Bayes* yaitu dengan penghitung probabilitas setiap kata yang terdapat pada SMS. Kata-kata pada SMS bersifat independen sehingga suatu kata tidak memiliki keterkaitan dengan kata lainnya.

7. *Cosine Similarity*

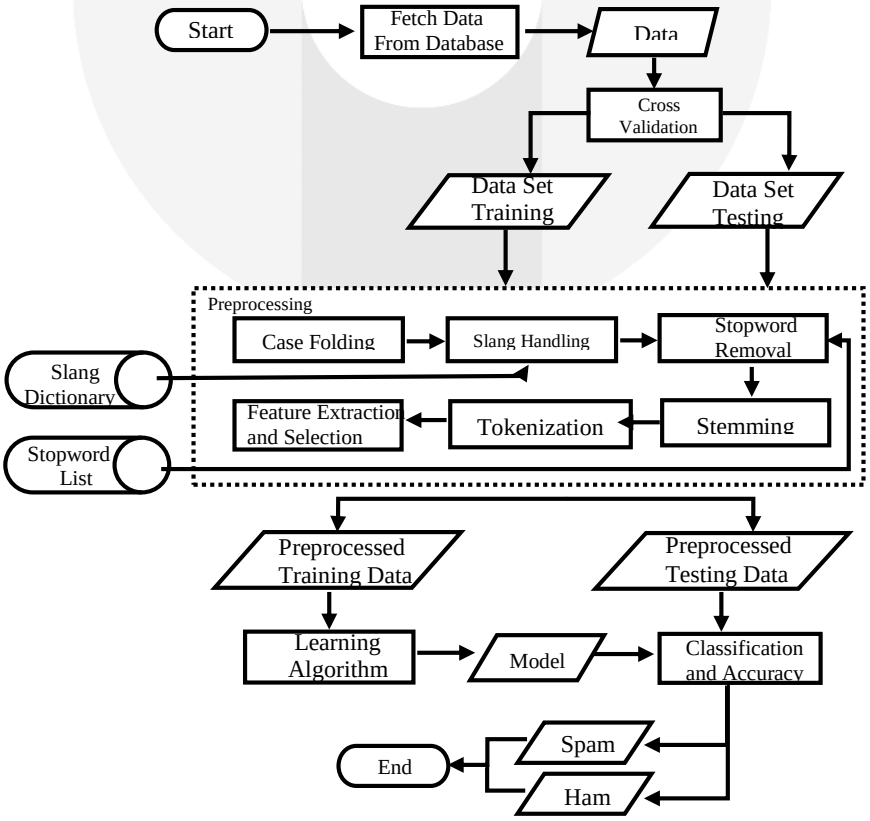
Penghitungan *cosine similarity* digunakan untuk melihat kemiripan suatu dokumen dengan dokumen-dokumen lainnya [2].

4. Perancangan dan Implementasi

4.1 Gambaran Umum Sistem

Pada tugas akhir ini dilakukan pengimplementasian sistem untuk mengklasifikasikan SMS menjadi *spam* atau *ham*. Metode yang digunakan adalah dengan *Collaborative Naïve Bayes* dan *Content-Based Naïve Bayes* secara terpisah.

Gambaran umum sistem yang dibagun pada Tugas Akhir ini ditunjukkan pada gambar 3-1.



Gambar 3-1 Gambaran umum sistem

Dari gambaran umum di atas, dapat diuraikan proses yang terdapat pada sistem:

1. Membangun *Database: Database* digunakan untuk menampung rekomendasi dari *user*. Dibangun sebanyak dua *database* yang mempunyai fungsi sebagai *database* latih dan database uji.
2. *Preprocessing*: Dilakukan *preprocessing* data pada data latih dan uji Content-Based Naïve Bayes sehingga diperoleh SMS yang lebih teratur dan memiliki informasi yang bagus. Sedangkan pada data testing Collaborative Naïve Bayes dilakukan preprocessing apabila terdapat data testing tidak memiliki rekomendasi dari beberapa user. Proses yang dilakukan adalah:
   a. *Case Folding*: *Case Folding* dilakukan untuk merubah semua huruf besar menjadi huruf kecil serta menghapus tanda baca.
   b. *Slang Handling*: S*lang Handling* dilakukan untuk memperbaiki beberapa kata yang disingkat atau kata-kata yang tidak lazim menjadi bentuk aslinya.
   c. *Stopword Removal*: *Stopword Removal* digunakan untuk menghilangkan kata-kata yang dianggap umum.
   d. *Stemming*: *Stemming* digunakan untuk mengubah semua kata yang memiliki imbuhan (affixes) dan menemukan kembali kata asalnya
   e. *Feature Extraction* dan *Selection*: *Feature Extraction* dan *Selection* digunakan untuk menemukan kata-kata yang tidak memiliki informasi yang berguna
   f. *Tokenization*: *Tokenization* digunakan untuk memecah kalimat menjadi kumpulan kata-kata.

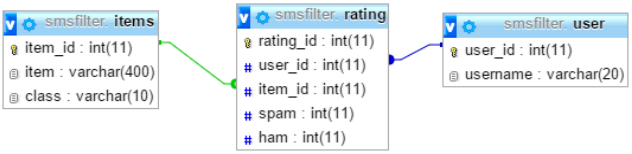
3. *Learning Algorithm*
   Proses *Learning Algorithm* digunakan untuk menerapkan metode *Collaborative* dan *Content-Based Naïve Bayes* yang menghasilkan model acuan pada saat melakukan *testing*.
4. Klasifikasi
   Data testing pada *Collaborative* akan diklasifikasikan menjadi kelas *spam* atau kelas *ham*. Proses pengklasifikasian ini didapatkan sesuai dengan model yang telah dibentuk sebelumnya pada proses *Learning Algorithm*.

4.2 Pengumpulan Rekomendasi *User*

Sebanyak 300 SMS dipilih secara acak dari kumpulan SMS pada UCI *Repository*. 300 SMS tersebut telah memiliki kelas masing-masing, yaitu *spam* atau *ham*. Dari total 300 SMS yang dipilih, terdapat 97 SMS *spam* dan 203 SMS *ham* dan keseluruhan dari 300 SMS tersebut disebar ke 42 responden yang diasumsikan sebagai *user* untuk mendapatkan rekomendasi. Terdapat tiga rekomendasi yang dapat diberikan oleh *user*, yaitu *Spam*, *Ham*, dan *Not Sure*. Rekomendasi yang diberikan oleh *user* tergantung pada kemampuan *user* untuk menganalisis SMS yang diberikan. Rekomendasi *Spam* diberikan ketika *user* melihat bahwa SMS yang diberikan adalah SMS *Spam* atau yang tidak diperlukan oleh *user* tersebut, begitu juga dengan rekomendasi *Ham*, yaitu dengan melihat bahwa SMS yang direkomendasikan adalah *Ham* atau SMS tersebut dibutuhkan oleh *user*. Rekomendasi *Not Sure* diberikan ketika *user* melihat bahwa SMS tersebut tidak merupakan SMS *Spam* atau *Ham* atau dapat dikatakan bahwa *user* mengabaikan SMS tersebut ketika SMS diklasifikasikan sebagai SMS *Spam* atau *Ham.*

4.3 Membangung *Database*

Database digunakan untuk menyimpan rekord dari setiap user yang berkontribusi memberikan rekomendasi terhadap 300 SMS. *Database* yang dibangun terdiri dari dua database yaitu *database training* dan *datbase testing*.



Gambar 3-2 Tabel Relasi

4.4 *Preprocessing* Data

Preprocessing terhadap data latih dilakukan sehingga data memiliki konten yang seragam, berguna, dan teratur sehingga memudahkan ketika dilakukan pemodelan dan mengeluarkan hasil yang lebih akurat dan mempercepat proses komputasi.

5. Pengujian dan Analisis

5.1 Tujuan Pengujian

Tujuan utama dalam melakukan pengujian sistem yang dibangun pada tugas akhir ini adalah untuk mengetahui kualitas dari pengklasifikasian SMS *Spam* yang telah dibangun.

Adapun pengujian yang akan dilakukukan terfokus pada:

1. Jumlah *user* dengan melihat perbedaan jumlah *user* yang memberikan rekomendasi, apakah terdapat perbedaan akurasi terhadap jumlah *user* yang berbeda.
2. Mengukur perbedaan nilai *F-Measure* dilihat dari perbedaan jumlah *user*.

5.2 Pembagian Data

Untuk mendapatkan data *training* dan data *testing* untuk pengujian pada tugas akhir ini, digunakan pembagian menggunakan teknik k-*fold cross validation* dengan membagi data menjadi k bagian dan mengkombinasikan hasil pembagian sebagai data latih dan data uji. Nilai k yang digunakan adalah 5 dan 10 untuk melihat pengaruh jumlah data terhadap hasil pengujian.

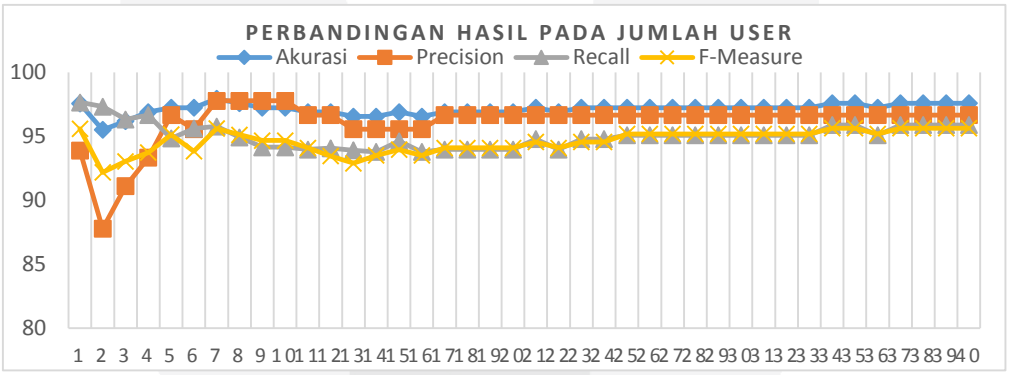
Dari kumpulan data yang telah didapatkan dari SMSSpamCollection (UCI *Repository*) diambil beberapa data percobaan untuk dilakukan survey pengambilan data rekomendasi kepada beberapa *user*. Data percobaan yang diambil sebanyak 300 data dengan rincian 97 *spam* dan 203 *ham*.

5.3 Skenario Pengujian

Skenario pengujian dilakukan dengan mengkombinasikan jumlah *user* terhadap pembagian *training* dan *testing* data 5-*fold* dan 10-*fold*. Jumlah *user* yang diukur dimulai dari 1 *user* sampai dengan 40 *user* secara berurutan.

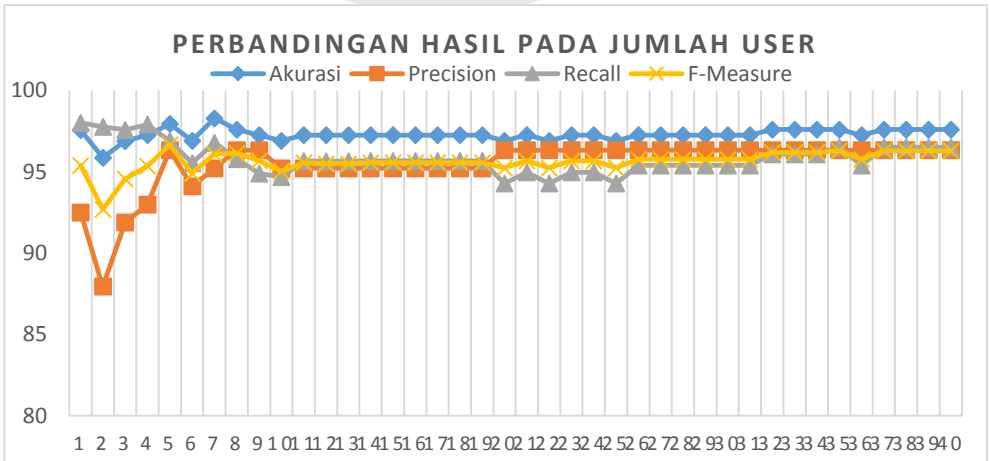
5.4 Hasil dan Analisis Pengujian

Setelah dilakukan pengujian untuk pembagian data 5-*fold* dan 10-*fold* dengan pengukuran jumlah user dan penggabungan data training pada Content-Based maka didapatkan hasil yang ditunjukkan pada gambar 4-1 dan 4-2.



Gambar 4-1 Hasil Pengujian Data 5-*fold*

Akurasi tertinggi muncul pada 25 *user* dan akurasi dapat dikatakan stabil pada 21 *user* yang memberikan rekomendasi. Didapatkan hasil rata-rata akurasi yaitu 97.12% dengan akurasi tertinggi didapatkan pada 7 *user* yaitu 97.93% dan akurasi terendah pada 2 *user* yaitu 95.51%.



Gambar 4-2 Hasil Pengujian Data 10-*fold*

Akurasi tertinggi muncul pada 25 *user* dan akurasi dapat dikatakan stabil pada user 11 *user* yang memberikan rekomendasi. Didapatkan hasil rata-rata akurasi yaitu 97.28% dengan akurasi tertinggi didapatkan pada 5 *user* yaitu 97.93% dan akurasi terendah pada 2 *user* yaitu 95.86%.

Dari pengujian yang dilakukan maka dapat disimpulkan:

1. Jumlah *user* yang melakukan rekomendasi mempengaruhi akurasi dari kelas suatu SMS.
2. Pengklasifikasian SMS berdasarkan konten dapat menambah keakurasian karena mendukung *Collaborative* ketika rekomendasi tidak terdapat pada SMS.
3. Rata-rata akurasi pada rekomendasi 40 *user* sebesar 97.12% untuk pembagian data 5-*fold* dan 97.28% untuk pembagian data 10-*fold*.
4. Rata-rata *precision* pada rekomendasi 40 *user* sebesar 96.12%, *recall* sebesar 95.04%, dan 94.61% untuk *F-Measure* pada 5-*fold*.
5. Rata-rata *precision* pada rekomendasi 40 *user* sebesar 95.45%, *recall* sebesar 95.81%, dan 95.62% untuk *F-Measure* pada 10 fold.
6. *F-Measure* meningkat untuk jumlah data latih yang banyak. Pada 5-*fold*, *F-Measure* yang didapatkan adalah sebesar 94.61% dan 95.62% untuk 10-*fold*.

Daftar Pustaka:

[1] C. Khemapatapan, "Thai-English Spam SMS Filtering", *16th Asia-Pacific Conference on Communications (APCC)*, 2010.

[2] F. Rahutomo, T. Kitasuka and M. Aritsugi, "Semanctic Cosine Similarity", *The 7th International Student Conference on Advanced Science and Technology ICAST*, 2012.

[3] H. Shirani-Mehr, "SMS Spam Detection using Machine Learning Approach".

[4] I. Ahmed, D. Guan and T. C. Chung, "SMS Classification Based on Naïve Bayes Classifier and Apriori Algorithm Frequent Itemset", *International Journal of Machine Learning and Computing*, Vol. 4, No. 2, 2014.

[5] K. Miyahara and M. J. Pazzani, "Collaborative Filtering with the Simple Bayesian Classifier", *PRICAI'00 Proceedings of the 6th Pacific Rim international conference on Artificial intelligence*, Pages 679-689.

[6] K. Yadav, P. Kumaraguru, A. Goyal, A. Gupta and V. Naik, "SMSAssassin: Crowdsourcing Driven Mobile-based System for SMS Spam Filtering", *Proceedings of the 12th Workshop on Mobile Computing Systems and Applications*, Pages 1-6, 2011.

[7] Kominfo, [Online]. Available: http://statistik.kominfo.go.id/site/data?idtree= 213&iddoc= 766&data-data_page=2. [Accessed 29 Oktober 2014].

[8] Menaka and Radha, "Text Classification using Keyword Extraction Technique", *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 3, Issue 12, 2013.

[9] RANKS NL, ranks.nl, [Online]. Available: http://www.ranks.nl/stopwords. [Accessed 29 March 2015].

[10] S. Berkovsky, P. Busetta, Y. Eytani, T. Kuflik and F. Ricci, "Collaborative Filtering over Distributed Environment".

[12] "Slang Dictionary - Text Slang & Internet Slang Words," NoSlang.com, 2005. [Online]. Available: http://www.noslang.com/dictionary/. [Accessed 23 June 2015].

[13] T. A. Almeida, J. M. Gomez and A. Yamakami, "Contributions to the study of SMS Spam Filtering: New Collection and Results", *Proceedings of the 11th ACM symposium on Document engineering*, Pages 259-262, 2011.

[14] T. B. Shahi and A. Yadav, "Mobile SMS Spam Filtering for Nepali Text Using Naïve", *International Journal of Intelligence Science*, pp.24-28, 2014.

[15] Tartarus [Online]. Available [http://tartarus.org/martin/PorterStemmer/def.txt] [Accessed 29 March 2015].

[16] UCI, "SMS Spam Collection Data Set," [Online]. Available: https://archive.ics.uci.edu/ml/datasets/SMS+Spam+Collection. [Accessed 23 June 2015].

[17] W. Deng and H. Peng, "Research On A Naïve Bayesian Based Short Message Filtering System", *Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian*, Pages 1233-1237, 2006.