

Klasifikasi Topik Multi Label pada Hadis Bukhari dalam Terjemahan Bahasa Indonesia Menggunakan Random Forest

Adhithia Wiraguna¹, Said Al Faraby, S.T., M.Sc², Prof. Dr. Adiwijaya, S.Si, M.Si³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹wiragynaadhithia@students.telkomuniversity.ac.id, ²saidal faraby@telkomuniversity.ac.id,

³adiwijaya@telkomuniversity.ac.id

Abstrak

Hadis merupakan hal yang wajib untuk dipelajari dan diamalkan oleh umat Islam. Terdapat banyak jenis ajaran yang dapat diambil oleh manusia dengan mempelajari hadis. Untuk membantu umat Islam dalam mempelajari hadis, dibutuhkan sistem klasifikasi multi label untuk mengategorikan Hadis Shahih Bukhari terjemahan bahasa Indonesia berdasarkan tiga topik yaitu larangan, anjuran dan informasi. Dalam membangun sistem klasifikasi teks, terdapat berbagai metode klasifikasi yang dapat digunakan, pada penelitian ini menggunakan *Random Forest* (RF). Kesederhanaan algoritma RF dan kemampuan yang baik dalam menghadapi data berdimensi tinggi, membuat RF merupakan metode yang cocok dalam melakukan klasifikasi teks. Namun belum banyak diketahui kemampuan RF untuk klasifikasi *multi label*. Penelitian ini menggunakan metode pendekatan *Problem Transformation* yaitu *Binary Relevance* (BR) dan *Label Powerset* (LP) untuk mengadaptasi RF dalam membangun sistem klasifikasi teks *multi label*. Hasil penelitian menunjukkan bahwa performansi *hamming loss* yang terbaik didapat dari sistem yang menggunakan BR dan tidak menggunakan *stemming* yaitu sebesar 0,0663. Hasil ini menunjukkan bahwa metode BR lebih baik daripada metode LP dalam mengadaptasi algoritma RF untuk melakukan klasifikasi *multi label* terhadap data hadis. Hal ini dikarenakan metode BR menghasilkan model klasifikasi sebanyak jumlah label pada data hadis dan pada sisi lainnya, hasil transformasi data dari penggunaan LP membuat data yang digunakan menjadi *imbalanced*.

Kata kunci : Klasifikasi, hadis, multi label, *random forest*, *problem transformation*

Abstract

Hadith is a mandatory thing to be studied and practiced by Muslims. There are many types of teachings that humans can take by studying the hadith. To assist Muslims in studying the hadith, a multi label classification system is needed to categorize Sahih Bukhari Hadi in Indonesian translation based on three topics, namely prohibition, advice and information. In building a text classification system, there are various classification methods that can be used, in this study using Random Forest (RF). The simplicity of the RF algorithm and good ability to deal with high dimensional data, make RF a suitable method of text classification. But, there is not widely known RF capability for the multi label classification. This study uses the Problem Transformation approach method, namely Binary Relevance (BR) and Label Powerset (LP) to adapt RF in building a multi label classification system. The results showed that the best hamming loss performance obtained from a system that used BR and does not use *stemming* which is equal to 0,0663. These results indicate that the BR method is better than the LP method in adapting the RF algorithm to perform multi label classification of hadith data. This is happened because the BR method produces a classification model of the number of labels in the hadith data and on the other hand, the transformation of data from the use of LP makes the data are imbalanced.

Keywords: Classification, hadith, multi label, random forest, problem transformation

1. Pendahuluan

1.1 Latar Belakang

Hadis yang merupakan dasar Agama Islam tentunya menjadikan hadis sebagai hal yang wajib untuk dipelajari dan diamalkan oleh umat Islam. Di dalam hadis, terdapat beberapa jenis ajaran yang dapat diambil oleh manusia. Beberapa hadis merupakan anjuran untuk umat Islam dalam menjalani kehidupan di dunia. Ada juga hadis yang berisikan larangan-larangan dalam berperilaku sebagai umat Islam. Namun, beberapa hadis ada yang bukan merupakan kedua hal tersebut, yang bisa dikatakan hanyalah sebagai informasi kepada umat Islam. Melihat hal ini, maka hadis dapat dibagi ke dalam tiga topik yaitu anjuran, larangan dan informasi. Untuk mempermudah umat Islam dalam mengamalkan kewajibannya yaitu mempelajari hadis, dibutuhkan sebuah sistem klasifikasi agar dapat mengategorikan hadis berdasarkan ketiga topik tersebut. Namun, ternyata secara makna, kategori suatu hadis juga dapat merupakan gabungan antara ketiga kategori tersebut yang berarti suatu hadis dapat memiliki lebih dari satu label. Model klasifikasi dengan model kasus tersebut disebut dengan klasifikasi *multi-label*.

Untuk membangun sistem klasifikasi teks, terdapat berbagai pendekatan klasifikasi yang dapat digunakan, salah satunya adalah *Random Forest* (RF). RF merupakan metode yang baik dalam klasifikasi teks karena

kesederhanaannya dan kemampuan generalisasinya dalam menghadapi data berdimensi tinggi [1]. Kemampuan RF dalam melakukan klasifikasi sudah terkenal baik untuk model klasifikasi *single label* [1] [2], namun belum banyak diketahui kemampuan RF dalam membangun model klasifikasi *multi label*.

Terdapat berbagai pendekatan untuk mengadaptasi permasalahan klasifikasi *multi label*, salah satunya merupakan *Problem Transformation Method*. *Problem Transformation Method* secara umum akan menyelesaikan model permasalahan *multi label* dengan mengubahnya menjadi model permasalahan *single label* dan mengintegrasikan hasilnya kembali ke dalam bentuk *multi label* seperti *Binary Relevance* dan *Label Powerset* [3].

Dengan begitu, permasalahan yang muncul adalah apakah algoritma Random Forest juga merupakan algoritma yang efisien dalam membangun sistem klasifikasi untuk mengategorikan data korpus hadis ke dalam tiga label tersebut dan apakah algoritma Random Forest dapat diadaptasi dengan baik untuk model permasalahan klasifikasi *multi label*. Oleh karena itu, fokus penelitian ini adalah pembangunan sebuah sistem klasifikasi *multi label* teks terhadap data Hadis Shahih Bukhari terjemahan Bahasa Indonesia ke dalam tiga label yaitu anjuran, larangan dan informasi dengan mengimplementasikan algoritma RF dan menganalisis pengaruh jenis metode *Problem Transformation* yang digunakan. Hasil klasifikasi yang didapat juga akan dianalisis untuk mendapatkan gambaran penyebab terjadinya kesalahan model dalam memprediksi kelas dan label dari data. Metode evaluasi yang digunakan dalam penelitian ini adalah *hamming loss*.

1.2 Tujuan

Penelitian yang dilakukan bertujuan untuk membangun sistem yang dapat melakukan klasifikasi topik dengan model klasifikasi *multi label* terhadap data Hadis Shahih Bukhari terjemahan Bahasa Indonesia ke dalam gabungan dari tiga label yaitu anjuran, larangan, dan informasi menggunakan *Random Forest* dan melakukan analisis performa dari sistem klasifikasi yang dibangun menggunakan metode evaluasi *hamming loss*.

Tabel 1. Tujuan dari penelitan yang dilakukan

No	Tujuan	Pengujian	Kesimpulan
1	Membangun sistem klasifikasi menggunakan <i>Random Forest</i> untuk melakukan klasifikasi topik pada data hadis terjemahan Bahasa Indonesia.	Sistem melakukan klasifikasi topik pada data hadis terjemahan Bahasa Indonesia.	Mendapatkan data hadis yang sudah berlabel.
2	Menganalisis performa sistem klasifikasi dalam melakukan klasifikasi data teks <i>multi label</i> .	Melakukan pengujian terhadap perbedaan penggunaan jumlah trees, jenis <i>Problem Transformation</i> yang digunakan, dan penggunaan <i>stemming</i> .	Mendapatkan nilai perfomansi yaitu <i>hamming loss</i> minimum dari sistem dalam pengujian.

2. Dasar Teori

2.1 Klasifikasi Hadis

Klasifikasi hadis telah pernah dilakukan dengan menggunakan empat belas kelas seperti pada penelitian [4]. Teks hadis juga telah dapat dikategorisasikan terhadap satu label antara kelas berupa anjuran, larangan dan informasi seperti pada penelitian [4][5]. Tetapi, penelitian-penelitian tersebut masih menggunakan model klasifikasi multi-class dan single-label.

Pada penelitian ini, akan dilakukan klasifikasi topik hadis dalam terjemahan bahasa Indonesia menggunakan tiga kelas berupa anjuran, larangan dan informasi. Model klasifikasi yang digunakan merupakan klasifikasi multi-label. Perbedaan dari penelitian-penelitian sebelumnya adalah dimana suatu hadis dapat dikategorisasikan ke dalam lebih dari satu label.

2.2 Klasifikasi *Multi Label*

Klasifikasi Multi-Label sudah menjadi perhatian para peneliti dalam mengembangkan dan menerapkannya di dunia Machine Learning. Pendekatan untuk menerapkan Multi-label pun sudah banyak diciptakan. Pendekatan dalam metode klasifikasi multi-label dibagi menjadi dua kategori utama, yaitu *Problem Transformation Methods* dan *Algorithm Adaptation Methods* [2].

Penelitian ini akan menggunakan metoda Random Forest (RF) yang merupakan pendekatan *Problem Transformation Methods*. Klasifikasi multi-label pada penelitian ini menggunakan tiga kelas yaitu berupa anjuran, larangan dan informasi.

2.3 TF-IDF

Ekstraksi fitur yang digunakan pada Tugas Akhir ini adalah Term Frequency Inverse Document Frequency (TF-IDF). Untuk menghitung TF-IDF, akan digunakan rumus seperti pada Persamaan 2.1.

$$w_{i,j} = tf \times idf = t f_{i,j} x \log \frac{D}{df} \quad (2.1)$$

Keterangan :

- $w_{i,j}$ = bobot kata t_j terhadap dokumen d_i
- $t f_{i,j}$ = jumlah kemunculan kata t_j dalam dokumen d_i
- D = jumlah dokumen
- df_i = jumlah kemunculan kata dalam dokumen D

2.4 Random Forest

Random Forest merupakan algoritma Ensemble Learning yang menggunakan dan membangun struktur Tree dalam tahapannya. Dalam penggunaannya, dibangun Decision Tree dengan memilih atau mengambil data secara acak. Untuk menentukan kelas suatu data, dalam Random Forest menggunakan sistem voting dari hasil berdasarkan Decision Tree tersebut[7] .

Berdasarkan teknis yang diterapkan oleh Random Forest, membuat hasil prediksi yang lebih efisien. Secara umum, Random Forest memiliki kelebihan antara lain [1]:

- Dapat mengatasi masalah over-fitting
- Tidak terlalu sensitif terhadap data outlier
- Parameter dapat diubah.

2.5 Hamming Loss

Untuk menentukan performansi dari penelitian yang dilakukan, diperlukan sebuah metoda untuk melakukan evaluasi. Pada penelitian atau Tugas Akhir ini, metoda evaluasi yang digunakan adalah evaluasi Hamming-Loss. Dalam metoda Hamming-Loss, perhitungan yang dilakukan adalah banyaknya kesalahan klasifikasi terhadap data yang diuji. Contoh jika seharusnya sebuah data hadis seharusnya diklasifikasikan sebagai label informasi dan anjuran, namun ketika diprediksi, data tersebut diklasifikasikan sebagai larangan. Hal-hal tersebut yang diperhitungkan dalam metoda Hamming-Loss. Performansi dalam metoda Hamming-Loss ditandai dengan representasi nilai dari $hloss(h)$. Semakin kecil nilai $hloss(h)$ maka berarti akurasi atau performansi dari sistem klasifikasi yang dibangun semakin tinggi atau semakin baik. Persamaan untuk menghitung $hloss(h)$ terlihat pada Persamaan 2.3

$$hloss(h) = \frac{1}{p} \sum_{i=1}^p \frac{1}{Q} |h(x_i) \Delta Y_i| \quad (2.2)$$

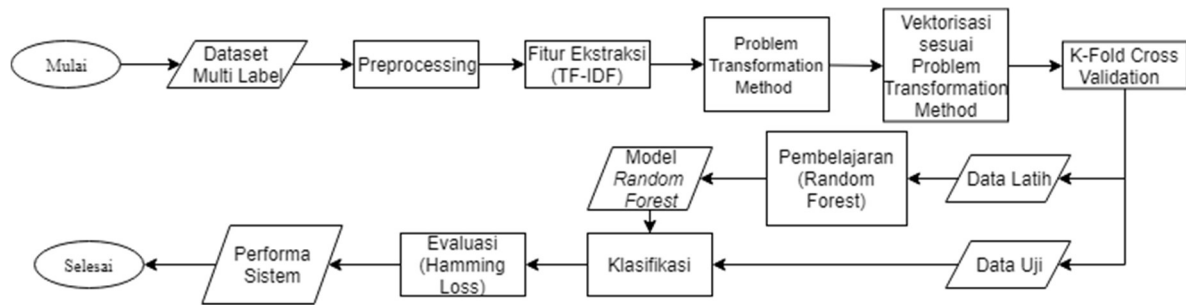
Keterangan:

- p = jumlah banyaknya data
- Q = jumlah banyaknya kelas
- $|h(x_i) \Delta Y_i|$ = jumlah banyaknya error atau kesalahan dalam klasifikasi yang terjadi

3. Sistem yang Dibangun

3.1 Gambaran Umum Sistem

Penelitian ini mengusulkan rancangan sistem yang dapat melakukan klasifikasi *multi label* topik terhadap data Hadis Shahih Bukhari terjemahan Bahasa Indonesia menggunakan algoritma *Random Forest*. Terdapat beberapa tahapan yang dilakukan dalam membangun rancangan sistem ini, yaitu *Problem Transformation*, *preprocessing*, ekstraksi fitur, model klasifikasi RF, dan evaluasi. Gambaran sistem secara umum dapat dilihat pada Gambar 1.



Gambar 1 Diagram Alur Sistem

Dataset yang digunakan dalam membangun rancangan sistem ini adalah hadis shahih Bukhari terjemahan Bahasa Indonesia berjumlah 2000 data. Data tersebut sebelumnya telah diberi label secara manual atau hand-labelling dimana terdiri dari tiga kelas yaitu anjuran, larangan dan informasi. Data hadis dapat merupakan lebih dari satu kelas.

Penelitian tugas akhir ini akan membangun sistem yang dibangun terdiri dari dua jenis, dibagi berdasarkan jenis metode *Problem Transformation* yang digunakan. *Problem Transformation* digunakan sebagai metode pendekatan yang membantu dalam membangun model klasifikasi multilabel. Metode-metode yang digunakan untuk *Problem Transformation* pada penelitian ini adalah *Binary Relevance* (BR) dan *Label Powerset* (LP).

Pada sistem yang dibangun menggunakan pendekatan BR, pembangunan model klasifikasi *multi label* nya adalah dengan membangun masing-masing model klasifikasi dari setiap label. Pada penelitian ini, terdapat tiga label yaitu anjuran, larangan dan informasi. Oleh karena itu, sistem dengan pendekatan BR akan membangun masing-masing model klasifikasi dari setiap label, yang mana berarti pada penelitian ini, akan dibangun tiga model klasifikasi untuk sistem pendekatan BR. Contoh data latih yang digunakan dalam membangun model klasifikasi dengan pendekatan BR terlihat pada Tabel 2.

Tabel 2 Representasi dataset dengan pendekatan *Binary Relevance* untuk label anjuran

Hadis	Anjuran
Jika salah seorang dari kalian berwudlu hendaklah dengan memasukkan air ke dalam hidung, barangsiapa beristinja' dengan batu hendaklah dengan bilangan ganjil. Dan jika salah seorang dari kalian bangun dari tidurnya, hendaklah membasuh kedua telapak tangannya sebelum memasukkannya dalam bejana air wudlunya, sebab salah seorang dari kalian tidak tahu ke mana tangannya bermalam.	1
Tumit-tumit yang tidak terkena air wudlu akan masuk ke dalam neraka. Beliau ucapkan itu hingga tiga kali.	0
Janganlah kalian berdusta terhadapku (atas namaku), karena barangsiapa berduasta terhadapku dia akan masuk neraka.	0

Sistem kedua yang dibangun dengan LP, akan mengubah data dengan model *multi label* menjadi model *single label*, yaitu dengan membuat label baru yang merupakan gabungan dari label pada model *multi label* sebelumnya. Dengan model data yang telah menjadi *single label* yang terlihat pada Tabel 3, data tersebut yang digunakan untuk dibangun model klasifikasi seperti layaknya pada model permasalahan klasifikasi *single label*.

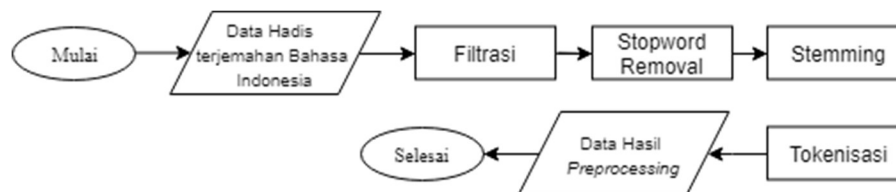
Tabel 3 Representasi dataset dengan pendekatan *Label Powerset*

Hadis	Kelas
Jika salah seorang dari kalian berwudlu hendaklah dengan memasukkan air ke dalam hidung, barangsiapa beristinja' dengan batu hendaklah dengan bilangan ganjil. Dan jika salah seorang dari kalian bangun dari tidurnya, hendaklah membasuh kedua telapak tangannya sebelum memasukkannya dalam bejana air wudlunya, sebab salah seorang dari kalian tidak tahu ke mana tangannya bermalam.	101
Tumit-tumit yang tidak terkena air wudlu akan masuk ke dalam neraka. Beliau ucapkan itu hingga tiga kali.	001

Janganlah kalian berdusta terhadapku (atas namaku), karena barangsiapa berdusta terhadapku dia akan masuk neraka.	011
---	-----

3.2 Preprocessing

Dataset yang digunakan tentunya masih berupa data mentah yang mana memiliki banyak noise. Sehingga diperlukan metode-metode pengolahan dataset untuk menghilangkan kata-kata yang tidak diperlukan atau yang berkemungkinan akan mengganggu dalam tahapan membangun sistem klasifikasi. Pada penelitian sistem ini, preprocessing memiliki tahapan tahapan seperti pada Gambar 2 . Keluaran dari tahapan preprocessing adalah data yang sudah bersih dari noise.

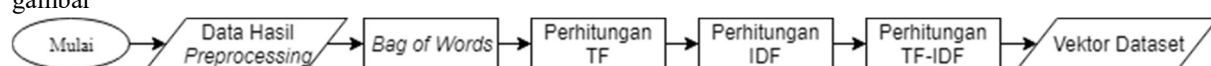


Gambar 2 Skema Preprocessing

- Filtrasi**
Pada proses filtrasi, tujuannya adalah menghilangkan Noise yang berupa tanda baca, angka dan karakter non-ASCII
- Stopword Removal**
Tujuan dari proses ini adalah menghilangkan kata-kata yang dianggap memiliki pengaruh kecil. Kata-kata yang dianggap kurang penting pengaruhnya dapat dilihat dari frekuensi kemunculannya yang jauh lebih banyak dibandingkan kata-kata lain. Implementasi metode *stopword removal* menggunakan bantuan *library* Sastrawi.
- Stemming**
Proses ini akan membuat seluruh kata dalam data masukan menjadi kata dasar dengan aturan tertentu. Implementasi metode *stemming* juga menggunakan bantuan *library* Sastrawi.
- Tokenisasi**
Dalam proses tokenisasi, setiap kata yang terdapat di dalam data masukan akan dipenggal yang mana akan membuat setiap kata menjadi satu nilai.

3.3 Ekstraksi Fitur (TF-IDF)

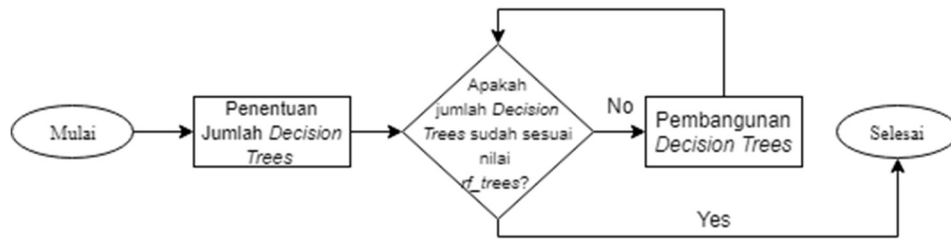
Ekstraksi fitur bertujuan untuk mengetahui bobot dari setiap fitur yang digunakan untuk tahapan pembelajaran sistem klasifikasi. Perhitungan bobot dari setiap fitur menggunakan TF-IDF. Bobot yang didapat adalah dengan perkalian antara TF dan IDF seperti pada Persamaan 2.1. Data yang digunakan untuk perhitungan bobot adalah data hasil *preprocessing*. Proses yang dilakukan pada tahapan ekstraksi fitur dapat dilihat pada gambar



Gambar 3 Tahapan Ekstraksi Fitur

3.4 Klasifikasi (Random Forest)

Data yang telah dilakukan proses ekstraksi fitur selanjutnya akan dibuat model klasifikasinya dengan metode Random Forest berdasarkan data latih. Metode Random Forest merupakan kumpulan dari sejumlah *decision tree* yang telah ditentukan dan dibangun dengan aturan aturan tertentu. Prediksi yang didapat adalah hasil proses voting dari pembelajaran berdasarkan seluruh *decision tree*. Pada model *random forest* yang dibangun, *decision trees* dibangun dengan aturan *entropy*. Gambaran umum dari pembentukan model klasifikasi *random forest* dapat dilihat pada Gambar 4.



Gambar 4 Skema pembangunan model klasifikasi *random forest*

4. Evaluasi

Setelah tahapan pembangunan sistem, akan dilakukan pengujian terhadap sistem untuk mendapatkan dan menganalisis performansi dari sistem. Penelitian tugas akhir ini telah dilakukan sebanyak 15 kali pengujian terhadap sistem.

4.1 Hasil Pengujian

Metode yang digunakan sebagai pendukung dalam pengujian sistem adalah metode evaluasi *hamming loss* dan metode *k-folds Cross Validation* dimana k yang digunakan pada penelitian ini adalah lima. Dalam pembangunan model klasifikasi RF, diambil seribu data latih dengan pengembalian untuk melatih setiap *tree*.

Tabel 4 Skenario 1 Pengujian Sistem

Parameter Pengujian			
Metode Problem Transformation	Fitur Ekstraksi	Rf_samples	Rf_trees
Binary Relevance	TF-IDF	1000	20,40,60,80,100
Label Powerset	TF-IDF	1000	20,40,60,80,100

Sesuai pada tabel 5, skenario pengujian tersebut dilakukan untuk mengetahui pengaruh dari jumlah *tree* (*rf_trees*) sebagai parameter dalam pembangunan model RF dan pengaruh perbedaan metode *Problem Transformation* dalam mengadaptasi RF untuk menyelesaikan permasalahan model klasifikasi *multi label*. Hasil performansi (*hamming loss*) dari pengujian sesuai dengan skema tabel 4, dapat dilihat pada tabel 5.

Tabel 5 Hasil Pengujian Skenario 1

Problem Transformation	Rf_trees	Hamming Loss					
		Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	avg
Binary Relevance	20	0,0775	0,06	0,0833	0,0675	0,0775	0,073167
	40	0,07583	0,06167	0,089167	0,063	0,079167	0,073833
	60	0,07083	0,063	0,08167	0,064167	0,079167	0,071833
	80	0,07167	0,0625	0,08167	0,064167	0,08167	0,072333
	100	0,08	0,0567	0,0875	0,0633	0,079167	0,073333
		Rata-rata					
Label Powerset	20	0,08833	0,06083	0,0867	0,07	0,0825	0,077667
	40	0,079167	0,059167	0,089167	0,065	0,0825	0,075
	60	0,078333	0,058333	0,08333	0,064167	0,084167	0,074667
	80	0,0775	0,0575	0,091667	0,0675	0,081667	0,075167
	100	0,079167	0,0575	0,089167	0,069167	0,083333	0,075667
		Rata-rata					

Tabel 6 Skenario 2 Pengujian Sistem

Parameter Pengujian				
Preprocessing	Metode Problem Transformation	Fitur Ekstraksi	Rf_samples	Rf_trees
Stemming	Binary Relevance	TF-IDF	1000	20
No Stemming	Binary Relevance	TF-IDF	1000	20
No Stemming	Binary Relevance	Rule-Based+TFIDF	1000	20

Skenario berikutnya yang dilakukan pada penelitian ini terlihat pada Tabel 6. Skenario pengujian tersebut dilakukan untuk mengetahui pengaruh penggunaan *stemming* dan pengaruh penggunaan antara TF-IDF dan metode *Rule-Based Extraction* berdasarkan penelitian [8] dalam melakukan ekstraksi fitur. Dengan *Rule-Based Extraction*, akan menghasilkan fitur yang berbeda dari *Bag of Words*. Pada *Bag of Words* semua kata unik pada dokumen teks akan digunakan sebagai fitur, namun pada *rule-based* hanya kata-kata yang sesuai dengan aturan yang sudah didefinisikan. Hasil performansi yang didapat dari skenario pengujian pada Tabel 6 dapat dilihat pada Tabel 7.

Tabel 7 Hasil Pengujian Skenario 2

No	Hamming Loss					
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	avg
1.	0,0775	0,06	0,0833	0,0675	0,0775	0,073167
2.	0,0683	0,04583	0,07583	0,0633	0,0783	0,0663
3.	0,07583	0,06	0,07917	0,06083	0,0875	0,07267

4.2 Analisis Pengujian

Observasi yang dilakukan untuk mengetahui pengaruh parameter jumlah *tree* terhadap performansi sistem adalah dengan melakukan klasifikasi menggunakan RF dengan rentang parameter *rf_trees* yaitu 20, 40, 60, 80 dan 100. Dapat dilihat pada tabel 5 bahwa nilai *hamming_loss* yang paling baik atau paling rendah didapat di saat *rf_trees* = 60 dan nilai performansi dari setiap rentang nilai parameter *rf_trees* yang digunakan adalah berbeda-beda. Hal ini disebabkan kemampuan *tree* yang dibangun akan berbeda-beda. Jika dalam RF yang dibangun didominasi dengan *tree* yang memiliki kemampuan generalisasi yang kuat maka akan semakin akurat hasil *voting* yang dihasilkan. Sehingga, pengujian ini membuktikan bahwa semakin besarnya jumlah *tree* yang digunakan pada pembangunan RF tidak selalu menghasilkan performansi yang lebih baik, karena faktor yang lebih menentukan adalah pada kemampuan generalisasi dari setiap *tree* yang dibangun.

Pada tahapan pengujian ini, observasi metode *Problem Transformation* dilakukan dengan membuat dan melakukan klasifikasi dengan tiga metode pendekatan *multi label* yang berbeda, yaitu BR dan LP. Dapat dilihat pada tabel 5, bahwa nilai rata-rata *hamming loss* yang didapat dari masing-masing pendekatan BR dan LP adalah 0,0728992 dan 0,0756336. Hal ini membuktikan bahwa performansi sistem yang dibangun menggunakan BR lebih baik dibandingkan dengan sistem yang menggunakan metode pendekatan LP.

Pada sistem yang menggunakan metode LP, model data yang awalnya *multi label* akan diubah ke bentuk data *single label* dan data *single label* ini yang akan digunakan untuk pembangunan model klasifikasi RF. Perubahan model data tersebut membuat kelas data akan menjadi lebih banyak dibandingkan dengan klasifikasi *multi label* dengan pendekatan BR

Tabel 8 Persebaran Kelas pada Sistem menggunakan metode LP

Kelas	Jumlah
001	1498
010	6
100	12
011	120
101	286
110	5
111	73
000	0

Persebaran kelas pada data yang sudah ditransformasi berdasarkan metode LP dapat dilihat pada Tabel 8. Berdasarkan Tabel 8, kelas '001' memiliki jumlah data yang paling banyak dan yang paling mendominasi pada dataset, sedangkan terlihat jumlah data yang kontras terhadap kelas yang lain, sehingga hal ini yang dinamakan dengan *imbalanced data* atau data yang tidak seimbang. *Imbalanced data* adalah disaat persebaran dari setiap kelas pada data yang digunakan yang tidak seimbang atau bahkan terdapat suatu kelas yang tidak memiliki representasi datanya. Hal ini yang menyebabkan kemampuan generalisasi atau klasifikasi dari model klasifikasi *multi label* yang menggunakan pendekatan LP tidak lebih baik dibandingkan klasifikasi *multi label* menggunakan pendekatan BR.

Tabel 9 Sampel Data yang Terklasifikasi Salah

Teks Hadis	Hasil Proses Teks Hadis	Target	Prediksi
Makanlah. Padahal saat itu mereka sedang berihram.	makan padahal itu sedang ihram	[1,0,1]	[0,0,1]
Makanlah. Sedang saat itu mereka sedang berihram.	makan sedang itu sedang ihram	[1,0,1]	[0,0,1]

Makanlah daging itu karena dia halal. Maka kemudian 'Amru berkata, kepada kami: Pergilah dan temui Shalih lalu tanyakan kepadanya tentang masalah ini dan masalah lainnya. Maka Shalih menemui kami lalu berkata, Bawalah kemari.	makan daging karena halal maka kemudian amru kata kami pergi temu shalih lalu tanya kepada masalah dan masalah lain maka shalih temu lalu kata bawa kemari	[1,0,1]	[0,0,1]
Kerjakanlah manasik dalam 'umrahmu seperti kamu kerjakan dalam hajimu.	kerja manasik umrah kamu kerja haji	[1,0,1]	[0,0,1]
Bersahurlah kalian, karena didalam sahur ada barakah	sahur kalian dalam sahur bara	[1,0,1]	[0,0,1]

Peneliti juga melihat dan melakukan analisis terhadap hasil klasifikasi yang didapat dari model klasifikasi yang dibangun. Peneliti melihat bahwa terdapat pola dalam hasil prediksi yang *error* atau tidak sesuai dengan target dari label data hadis. Hasil klasifikasi dari sistem yang dibangun dengan metode pendekatan BR, dari 66 data yang terprediksi salah, terdapat 29 data yang jenis label seharusnya terprediksi sebagai Anjuran dan Informasi, tetapi sistem hanya memprediksi label dari data tersebut sebagai Informasi.

Hal tersebut terjadi karena data hadis yang dihasilkan setelah dilakukan *preprocessing*, dengan kasat mata saja sulit untuk menentukan jenis topik dari hadis tersebut merupakan anjuran atau tidak. Data hadis yang merupakan anjuran, secara umum memiliki kata yang bersifat perintah, yang pada umumnya kata tersebut berakhiran *-lah*. Namun, ciri utama tersebut hilang setelah dilakukan salah satu metode *preprocessing* yaitu *stemming* terhadap data, seperti yang terlihat pada Tabel 10.

Dari hasil analisa tersebut, peneliti melakukan pengujian terhadap penggunaan *stemming*. Hasil pada Tabel 7 membuktikan bahwa hasil performansi yang didapat dari model yang dibangun tanpa *stemming* lebih baik dibandingkan dengan menggunakan *stemming*. Namun dengan menghilangkan proses *stemming*, waktu yang diperlukan untuk proses klasifikasi akan lebih lama karena fitur yang diproses semakin lama, sehingga peneliti mencoba untuk melakukan pengujian dengan penambahan metode ekstraksi fitur yaitu *rule-based* [8] dan pembobotan TF-IDF. Dari ketiga skenario tersebut, hasil performansi yang paling baik adalah model dengan pembangunan seluruh fitur tanpa *stemming* dan menggunakan TF-IDF yaitu **0,0663**. Hal ini dikarenakan *stemming* dapat mengubah makna dari kata dan *rule-based* belum lebih baik dan lengkap dalam penentuan aturan untuk ekstraksi fitur.

Tabel 10 Sampel Dataset

No	Teks Hadis	Target
1	Turunlah disini dan siapkan minuman buatku. Orang itu berkata: Wahai Rasulullah, bukankah masih ada matahari? Beliau berkata, lagi: Turunlah (berhenti disini) dan siapkan minuman buatku. Orang itu berkata, lagi: Wahai Rasulullah, bukankah masih ada matahari? Beliau berkata, lagi: Turunlah dan siapkan minuman buatku. Maka orang itu berhenti lalu memberikan minuman kepada Beliau, lalu Beliau minum kemudian melempar sesuatu dengan tangan Beliau ke suatu arah lalu bersabda: Apabila kalian telah melihat malam sudah datang dari arah sana maka orang yang puasa sudah boleh berbuka. Hadits ini dikuatkan pula oleh Jarir dan Abu Bakar bin 'Ayyasy dari Asy-Syaibaniy dari Ibnu Abu Awfa Radhiyallahu'anhu berkata: Aku pernah bersama Nabi dalam suatu perjalanan.	[0,0,1]
2	Jika kamu berkata kepada temanmu pada hari Jum'at 'diamlah', padahal Imam sedang memberikan khutbah maka sungguh kamu sudah berbuat sia-sia (tidak mendapat pahala).	[0,0,1]
3	Pertama kali yang akan kita kerjakan pada hari ini adalah shalat, kemudian kembali pulang dan menyembelih hewan kurban. Maka barangsiapa mengerjakan seperti ini berarti dia telah memenuhi sunnah kami.	[1,0,1]
4	Shalat Hari Raya 'Idul Fitri dua rakaat dan tidak shalat sebelum atau sesudahnya. Kemudian beliau mendatangi para wanita dan memerintahkan mereka untuk bersedekah. Maka para wanita memberikan sedekah hingga ada seorang wanita yang memberikan anting dan kalungnya.	[1,0,1]

Pada pembangunan sistem tanpa menggunakan *stemming* walaupun meningkatkan rata-rata performansi dari model ternyata tetap menghasilkan prediksi yang sama pada data Tabel 9. Lalu dilakukan peninjauan kembali terhadap data yang terklasifikasi salah dengan penggunaan sistem tanpa menggunakan *stemming*. Didapatkan gambaran bahwa sistem akan sering melakukan kesalahan pada klasifikasi pada label anjuran. Pada Tabel 10 yang merupakan sampel dari data yang telah dilakukan pelabelan secara manual sebelumnya, terlihat bahwa banyak dataset yang digunakan masih ambigu jika dilihat secara makna per kata. Contoh pada data nomor 2 pada Tabel 10, jika dilihat secara makna per kata, seharusnya data tersebut juga berkemungkinan untuk dikatakan sebagai hadis Anjuran namun pada dataset yang digunakan, data tersebut tidak dianggap sebagai Anjuran. Sebaliknya pada

data nomor 3 yang tidak terlalu terlihat sebagai hadis Anjuran namun didefinisikan sebagai hadis Anjuran pada dataset yang digunakan.

5. Kesimpulan

Berdasarkan hasil pengujian dan analisis yang telah dilakukan maka dapat diambil kesimpulan sebagai berikut:

1. Performansi *hamming loss* yang dihasilkan dari sistem yang dibangun dengan penggunaan nilai parameter jumlah *tree* dengan rentang nilai yaitu 20, 40, 60, 80, dan 100 adalah berbeda-beda. Hal ini terjadi karena setiap pembangunan *tree* pada RF berdasar pada data latih yang diambil secara acak dan dengan pengembalian, sehingga semakin besarnya jumlah *tree* yang digunakan pada pembangunan RF tidak selalu menghasilkan performansi yang lebih baik.
2. Sistem yang dibangun menggunakan RF dengan menggunakan model pendekatan *multi label* BR lebih baik dibandingkan sistem yang dibangun menggunakan pendekatan LP dan LSP, yaitu rata-rata performansi *hamming loss* yang didapat untuk sistem yang menggunakan pendekatan BR adalah 0,0728992 sedangkan hasil yang didapat pada sistem yang menggunakan pendekatan LP adalah 0,0756336. Hal tersebut terjadi karena perubahan model data yang dilakukan pada LP membuat persebaran kelas data yang digunakan untuk membangun model klasifikasi RF menjadi *imbalanced*.
3. Hasil dari adaptasi model klasifikasi RF ke bentuk klasifikasi *multi label* dengan pendekatan *Problem Transformation* cukup baik. Dari beberapa skema pengujian, pada sistem yang menggunakan metode pendekatan *Binary Relevance* dan tanpa *stemming* menghasilkan performansi terbaik dengan nilai *hamming loss* sebesar 0,0663.
4. Diperlukan juga perbaikan pada dataset yang digunakan. Sebaiknya pelabelan data dilakukan oleh ahli dalam keilmuan mengenai Hadis dan persebaran data.

Daftar Pustaka

- [1] B. Xu, X. Guo, Y. Ye, and J. Cheng, "An improved random forest classifier for text categorization," *J. Comput.*, vol. 7, no. 12, pp. 2913–2920, 2012.
- [2] et al. Liu, Feng, "MLRF: multi-label classification through random forest with label-set partition.," *Int. Conf. Intell. Comput.*, 2015.
- [3] G. Tsoumakas and I. Katakis, "Multi-Label Classification," *Int. J. Data Warehous. Min.*, vol. 3, no. 3, pp. 1–13, 2007.
- [4] F. Harrag, E. El-Qawasmeh, and P. Pichappan, "Improving Arabic text categorization using decision trees," *2009 1st Int. Conf. Networked Digit. Technol. NDT 2009*, pp. 110–115, 2009.
- [5] E. Riviera, R. Jasin, and S. Al-faraby, "Klasifikasi Anjuran, Larangan dan Informasi pada Hadis Sahih Al-Bukhari berdasarkan Model Unigram menggunakan Artificial Neural Network (ANN)," vol. 4, no. 3, pp. 4683–4694, 2017.
- [6] D. T. M. Syair Audi Liri Sacra, Said Al Faraby, "Bukhari Menggunakan Naïve Bayes Classifier Classification of Suggestion, Prohibition and Information of Shahih Bukhari Hadith Using Naïve Bayes Classifier," vol. 4, no. 3, pp. 4794–4802, 2017.
- [7] J. Ali, R. Khan, N. Ahmad, and I. Maqsood, "Random forests and decision trees," *IJCSI Int. J. Comput. Sci. Issues*, vol. 9, no. 5, pp. 272–278, 2012.
- [8] G. Mediamer, "Pengembangan Ekstraksi Fitur Berbasis Aturan dalam Klasifikasi Teks Multilabel untuk Hadits Bukhari dalam Terjemahan Bahasa Indonesia Program Studi Sarjana S1 Teknik Informatika Fakultas Informatika Universitas Telkom Bandung," 2018.
- [9] Asriyanti Indah Pratiwi, Adiwijaya. 2018. On the Feature Selection and Classification Based on Information Gain for Document Sentiment Analysis. Applied Computational Intelligence and Soft Computing, 2018.
- [10] Mubarok, M.S., Adiwijaya and Aldhi, M.D., 2017. Aspect-based sentiment analysis to review products using Naïve Bayes. In AIP Conference Proceedings (Vol. 1867, No. 1, p. 020060). AIP Publishing
- [11] Aziz, R.A., Mubarok, M.S. and Adiwijaya, A., 2016, September. Klasifikasi Topik pada Lirik Lagu dengan Metode Multinomial Naive Bayes. In Indonesia Symposium on Computing (IndoSC) 2016
- [12] Reynaldi Ananda Pane, Mohamad Syahrul Mubarok, Nanang Saiful Huda, Adiwijaya, 2018. A Multi-label Classification on Topics of Quranic Verses in English Translation using Multinomial Naive Bayes. In 2018 6th International Conference on Information and Communication Technology (ICoICT). IEEE
- [13] Al Mira Khonsa Izzaty, Mohamad Syahrul Mubarok, Nanang Saiful Huda, Adiwijaya, 2018. A Multi-label Classification on Topics of Quranic Verses in English Translation Using Tree Augmented Naïve Bayes. In 2018 6th International Conference on Information and Communication Technology (ICoICT).

- [14] Abubakar, M.Y., Adiwijaya, and Adiwijaya, 2018. Multi-Label Topic Classification for Hadith of Bukhari in Indonesian Translation using Information Gain and Backpropagation Neural Network. In 2018 International Conference on Asian Language Processing (IALP). IEEE.IEEE.