

Implicit Social Trust Dan Support Vector Regression Untuk Sistem Rekomendasi Berita

Implicit Social Trust and Support Vector Regression for News Recommender System

¹Melita Widya Ningrum dan ²Wijanarto

Program Studi Ilmu Komputer, Universitas Dian Nuswantoro, Semarang
e-mail: ¹111201307651@mhs.dinus.ac.id, ²wijanarto@dsn.dinus.ac.id

Abstrak

Situs berita merupakan salah satu situs yang sering diakses masyarakat karena kemampuannya dalam menyajikan informasi terkini dari berbagai topik seperti olahraga, bisnis, politik, teknologi, kesehatan dan hiburan. Masyarakat dapat mencari dan melihat berita yang sedang populer dari seluruh dunia. Di sisi lain, melimpahnya artikel berita yang tersedia dapat menyulitkan pengguna dalam menemukan artikel berita yang sesuai dengan ketertarikannya. Pemilihan artikel berita yang ditampilkan ke halaman utama pengguna menjadi penting karena dapat meningkatkan minat pengguna untuk membaca artikel berita dari situs tersebut. Selain itu, pemilihan artikel berita yang sesuai dapat meminimalisir terjadinya banjir informasi yang tidak relevan. Dalam pemilihan artikel berita dibutuhkan sistem rekomendasi yang memiliki pengetahuan mengenai ketertarikan atau relevansi pengguna akan topik berita tertentu. Pada penelitian ini, peneliti membuat sistem rekomendasi artikel berita pada *New York Times* berbasis *implicit social trust*. *Social trust* dihasilkan dari interaksi antara pengguna dengan teman-temannya dan bobot kepercayaan teman pengguna pada media sosial Twitter. Data yang diambil merupakan data pengguna Twitter, teman dan jumlah interaksi antar pengguna berupa *retweet*. Sistem ini memanfaatkan algoritma *Support Vector Regression* untuk memberikan estimasi penilaian pengguna terhadap suatu topik tertentu. Hasil pengolahan data dengan *Support Vector Regression* menunjukkan tingkat akurasi dengan *MAPE* sebesar 0,8243075902233644%.

Keywords : *Twitter, Rekomendasi Berita, Social Trust, Support Vector Regression*

Abstract

News site is one of the most frequently accessed sites because it can provide actual information from various topic such as sport, business, politic, technology, health and entertainment. The overflow of article can be difficult for users to find article that relevance to them. Choice of article that shows up on a user's home page become an important task because it can raise user's will to read article from the current site. Recommender system is needed when choicing article for a user. This system must have base knowledge of the user whether his interest or relevance for spesific news topic. In this research, a recommender system will be built. This system focus on giving article recommendation from *New York Times*. *Social trust* link will be implemented as a base of the system. *Social trust* created from interaction between a user and his friend also trust value from each friend in social media Twitter. Data is from Twitter including user's, friend's and amount of interaction between users called *retweet*. This

system use Support Vector Regression algorithm to estimate user's rating to specific news topic. The result shows accuracy of this system with MAPE 0,8243075902233644%.

Keywords : Twitter, News Recommender System, Social Trust, Support Vector Regression

1. PENDAHULUAN

Membaca berita secara *online* merupakan salah satu cara yang populer untuk membaca artikel berita saat ini[1]. Berita *online* dipilih karena keterbacaan dan akses yang lebih mudah daripada berita di surat kabar[2]. Selain itu, artikel berita yang dapat ditampung oleh situs berita lebih banyak dan beragam, mencakup belahan dunia manapun karena semua terhubung dengan internet.

Di sisi lain, melimpahnya artikel berita dapat menyulitkan pembaca untuk menemukan berita yang relevan atau sesuai dengan ketertarikannya[2]. Penting bagi penyedia situs berita untuk membantu pengguna dalam menyediakan artikel yang menarik sesuai dengan minat pengguna[1]. CNN.com berintegrasi dengan Facebook's "social toolkit" untuk mengatasi ketidaksesuaian berita yang ditampilkan dengan keinginan pengguna. Integrasi ini ditujukan untuk memberikan pengalaman berita personal yaitu dengan menampilkan artikel yang direkomendasikan dan disebar oleh teman pengguna[3].

Banyak penelitian dilakukan dengan memanfaatkan *real-time social networking sites* seperti Facebook, Google Plus dan Twitter. Penelitian ini ditujukan untuk mengidentifikasi berita yang paling populer atau berita terbaru [2] seperti yang dilakukan oleh Jonnalagedda[2] yang memanfaatkan Twitter untuk membangun sistem rekomendasi berita personal. Sistem ini menggunakan metode *hybrid* antara *Popularity Based News Recommender* dan *Profile Based News Recommender*. Twitter digunakan karena terdapat fitur tweet di mana pengguna dapat menyalurkan opini mereka hanya dalam 140 karakter sehingga tweet lebih mudah diproses. Alahmadi[4] juga memanfaatkan Twitter API pada Java untuk mempersiapkan dataset pada penelitiannya yang memanfaatkan *implicit social trust* dari pengguna twitter sebagai basis sistem rekomendasi dengan algoritma genetika untuk optimasi *social trust parameters*. Selain Alahmadi, Chaochao Chen [5] juga menerapkan *social trust relationships* pada sistem rekomendasi yang dibangunnya. Sistem rekomendasi ini menitikberatkan pada pengaruh hubungan sosial antar pengguna yang dapat mempengaruhi kegemaran pengguna. Ide ini dicetuskan karena sistem rekomendasi tradisional tidak memberikan hasil yang realistis[5].

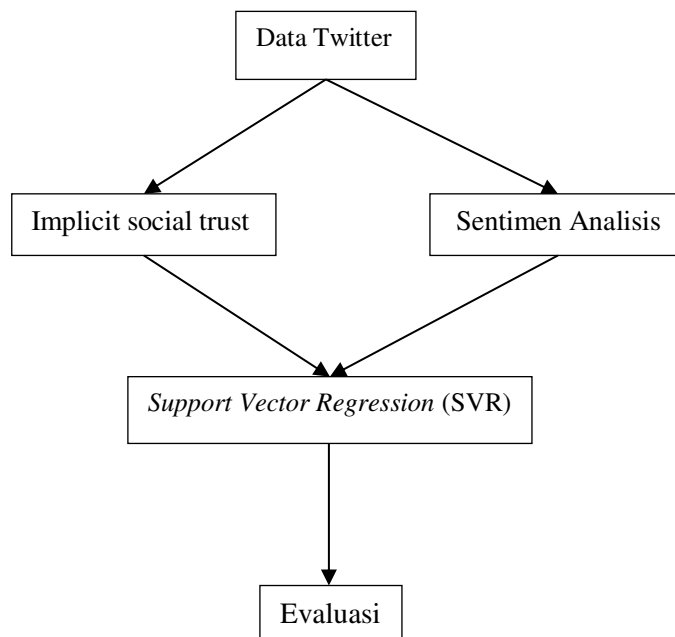
Menjawab permasalahan pengguna dalam memilah informasi dari artikel berita yang relevan maka sistem rekomendasi berita dapat menjadi solusi. Sistem rekomendasi akan mempelajari aktivitas pengguna di media sosial sehingga dapat memberikan daftar rekomendasi berita sesuai dengan preferensi pengguna.

Penelitian ini memanfaatkan *implicit social trust* pengguna di media sosial dan algoritma Support Vector Regression dalam estimasi rating. Support Vector Regression dipilih karena memiliki tingkat akurasi hasil yang lebih baik dibandingkan dengan *linear regression* dan *random forest*. Hal ini telah dibuktikan pada kasus sistem rekomendasi film di mana Support Vector Regression memberikan hasil Mean Absolute Percentage Error 1% di bawah algoritma lainnya[4]. Kedepan hasil penelitian ini dapat di pakai sebagai perbandingan atau bahan pengambilan keputusan oleh otoritas yang berkepentingan dalam menyikapi peran media sosial dan tingkat kepercayaan masyarakat terhadap suatu permasalahan tertentu.

2. METODE PENELITIAN

Pada penelitian ini, penyelesaian masalah dilakukan dengan beberapa tahap yang direpresentasikan pada Gambar 1. Langkah pertama adalah pengambilan data Twitter dengan memanfaatkan Twitter API. Untuk data berita diambil dari New York Times dengan NYT API. Data dari Twitter terdiri dari *tweet*, *username*, jumlah *retweet* dengan semua teman pengguna,

jumlah *follower*, jumlah *following* dan jumlah *retweet* antara pengguna dengan masing-masing teman. Langkah kedua merupakan pengolahan data yang terdiri dari dua tahap yaitu ekstraksi *implicit social trust* dan sentimen analisis. Data dari Twitter berupa *username*, jumlah *retweet* dengan semua teman pengguna, jumlah *follower*, jumlah *following* dan jumlah *retweet* antara pengguna dengan masing-masing teman diolah untuk diambil nilai *implicit social trust* pada jejaring pertemanan pengguna. Data berupa *tweet* diolah dengan sentimen analisis sehingga menghasilkan rating pengguna akan suatu topik berita tertentu.



Gambar 1 Metode Penelitian

Berikut langkah kedua yang dilakukan pada penelitian:

1. Ekstraksi *implicit social trust* [4]

$$RT = \frac{RT_{u,f}}{RT_{u,F}} , \dots \dots \dots (1)$$

Di mana $RT_{u,f}$ adalah jumlah retweet yang dilakukan oleh pengguna u terhadap temannya f . $RT_{u,F}$ adalah jumlah retweet yang dilakukan oleh pengguna u ke semua temannya F .

Parameter berupa kecenderungan kepercayaan terhadap teman f yang disimbolkan dengan L sebagai berikut:

$$L = \frac{Follower}{Follower+Following} , \dots \dots \dots (2)$$

Di mana $Follower$ merupakan jumlah *follower* dari teman f sedangkan $Following$ merupakan jumlah *following* dari teman f .

Berdasarkan RT dan L dapat ditemukan hubungan kepercayaan antara pengguna u dengan temannya f sesuai dengan rumus:

$$trust_{u,f} = \frac{1}{2}(RT + L) ,.....(3)$$

2. Sentimen Analisis

a. Tokenisasi

Proses memecah sebuah kalimat atau *stream* dari karakter menjadi *tokens* atau kata. Pada proses ini beberapa karakter akan dihilangkan seperti tanda baca titik koma “,” dan spasi[6].

b. Filtering

Proses menghilangkan kata-kata yang tidak berpengaruh atau *stopwords* pada artikel seperti “a” dan “the” serta “#”.

c. Stemming

Proses mengembalikan kata dasar dari suatu token. Apabila token tidak ditemukan dalam library maka akan mengembalikan token tersebut.

d. Term Weighting

Proses menentukan bobot dari tiap kata. Untuk pembobotan yang bertujuan menganalisa sentimen opini dapat memanfaatkan SentiWordNet. Dengan SentiWordNet dapat diambil nilai positif, negatif serta objektif dari suatu kata. Nilai objektif didapatkan sesuai dengan rumus berikut[7]:

$$NilaiObjektif = 1 - (NilaiPositif + NilaiNegatif),.....(4)$$

e. Perhitungan rating

Dengan memanfaatkan hasil perhitungan bobot dari SentiWordNet, term dapat dikelompokkan menjadi 5 kelas sesuai dengan penelitian yang dilakukan oleh Alahmadi[4].

$$wc_j = \begin{cases} 5, & \text{jika StatusPositif dan senti}_{wc_j} \geq 0.5 \\ 4, & \text{jika StatusPositif dan senti}_{wc_j} < 0.5 \\ 2, & \text{jika StatusNegatif dan senti}_{wc_j} < 0.5 \\ 1, & \text{jika StatusPositif dan senti}_{wc_j} \geq 0.5 \\ 3, & \text{lainnya} \end{cases} ,.....(5)$$

Untuk memprediksi sentiment rating dari opini teman digunakan rumus sebagai berikut:

$$SR(mr) = \sum c_j \in C C_j * E'_{mr,c_j},.....(6)$$

di mana c_j adalah label kelas numerik dari j di set C . E'_{mr,c_j} merupakan probabilitas bahwa mr memiliki sebuah rating c_j . E'_{mr,c_j} didapatkan sesuai dengan rumus berikut:

$$E'_{mr,c_j} = \frac{E(mr,c_j)}{\sum c_i \in E_{mr,c_i}},.....(7)$$

$E(mr,c_j)$ merupakan nilai yang diestimasi untuk mr berdasarkan wc_j dengan rumus sebagai berikut:

$$E(mr,c_j) = P(c_j)SI_{mr,c_j},.....(8)$$

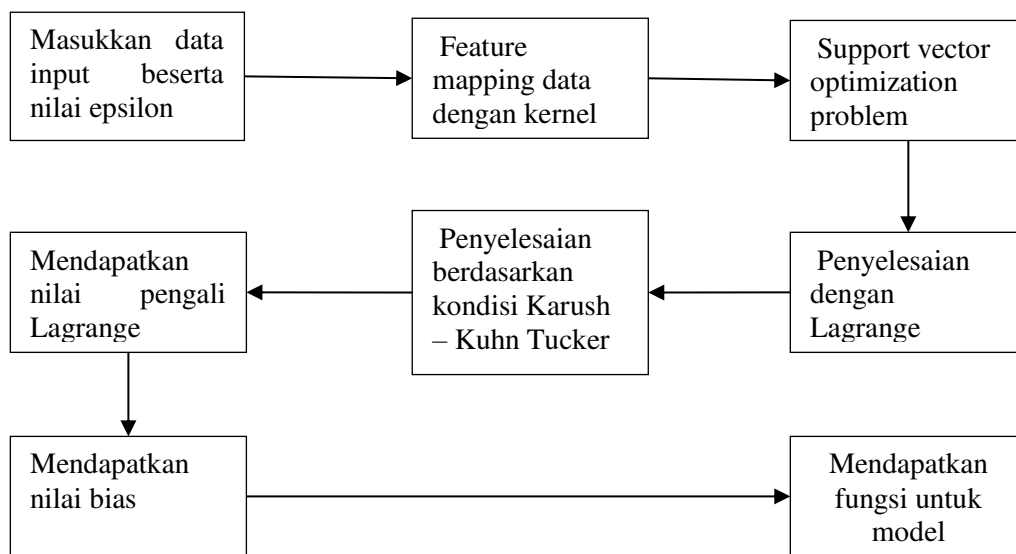
$P(c_j)$ merupakan probabilitas prior dari c_j pada dataset.

$$P(c_j) = \frac{N(c_j)}{N},.....(9)$$

di mana $N(c_j)$ adalah jumlah contoh yang termasuk ke dalam kelas c_j dan N adalah jumlah contoh data.

$$SI(mr, c_j) = w_{cj}(senti_{wjc}), \dots \dots \dots (10)$$

Data hasil pengolahan *implicit social trust* dan sentimen analisis kemudian digabungkan untuk menjadi *input* LIBSVM[8]. Data tersebut kemudian dibagi menjadi data latih dan data uji di mana data latih dimanfaatkan dalam proses pemodelan dengan algoritma *Support Vector Regression* (SVR) pada langkah keempat. Berikut representasi dari langkah-langkah Algoritma *Support Vector Regression* untuk lebih lengkap penjelasannya dapat dilihat pada *A tutorial on support vector regression*[9], dalam paper ini coba di sarikan dalam bahasa penulis supaya lebih sederhana :



Gambar 2 Blok Diagram Support Vector Regression (SVR) [9]

Setelah model dihasilkan, dilakukan langkah kelima yaitu evaluasi dengan pengujian menggunakan data uji dan *Mean Absolute Percentage Error* (MAPE) [10] sehingga didapatkan nilai akurasi hasil rekomendasi dengan *Support Vector Regression*.

$$MAPE = \frac{100}{N} \sum_{u \in N} \left| \frac{r_{u,i} - \hat{r}_{u,i}}{r_{u,i}} \right|, \dots \dots \dots (11)$$

Di mana $r_{u,i}$ adalah nilai target benar sedangkan $\hat{r}_{u,i}$ adalah nilai target prediksi. N merupakan jumlah dari data uji.

3. HASIL DAN PEMBAHASAN

Data pengguna yang diolah dibagi menjadi data *training* sebanyak 30 pengguna dalam blok warna terang dengan huruf normal pada tabel 1 dan data *uji* sebanyak 6 pengguna dalam blok warna gelap dengan huruf tebal pada tabel 1. Data yang diambil terdiri dari *username*, *tweet*, jumlah *retweet*, jumlah *follower* dan jumlah *following* yang sampelnya direpresentasikan pada Tabel 1.

Kolom 1 berisi *username* 36 pengguna. Kolom 2 berisi jumlah *retweet* yang dilakukan oleh ke-36 pengguna dengan seluruh teman. Kolom 3 menunjukkan jumlah *tweet* yang dilakukan oleh pengguna yang terdiri dari 3 *tweet*. Kolom 4 berisi topik dari ketiga *tweet* pengguna. Topik *Politic* disimbolkan dengan P, *Business* disimbolkan dengan B, *Tech* disimbolkan dengan T, *Health* disimbolkan dengan H, *Sport* disimbolkan dengan S. Kolom 5 berisi jumlah teman dari pengguna. Kolom 6 berisi jumlah teman dari pengguna. Kolom 7 berisi topik *tweet* yang dilakukan oleh masing-masing teman. Kolom 8 berisi jumlah *retweet* dari pengguna dengan kelima temannya. Kolom 9 berisi jumlah *follower* masing-masing teman dari pengguna. Kolom 10 berisi jumlah *following* masing-masing teman dari pengguna. Pengolahan data pengguna untuk diambil nilai *social trust* direpresentasikan pada Algoritma 1.

Tabel 1. Data Twitter

No	1	2	3	4	5	6	7	8	9
1	leoboudo	75	3	(2 A, 1 P)	5	(3 P, 1 H, 1 A)	(1, 1, 2, 2, 2)	(730607, 1232, 3968, 464758, 160001)	(891, 977, 640, 1293, 991)
2	Chrisness88	14	3	(2 A, 1 B)	5	(5 P)	(1, 1, 2, 1, 1)	(221, 3030, 451920, 118, 75443)	(328, 621, 419, 524, 13)
3	PinedoPR	12 1	3	(1 A, 1 H, 1 P)	5	(5 T)	(5, 5, 3, 4, 8)	(3579, 977584, 533, 323, 722)	(679, 275, 1378, 243, 675)
4	mamarocks54	61	3	(1 A, 1 H, 1 P)	5	(3 P, 1 B, 1 T)	(5, 5, 2, 2, 1)	(44958, 2077981, 1, 2870, 3771230)	(2222, 747, 1, 4759, 938)
5	msresister	12 8	3	(1 A, 1 P, 1 B)	5	(4 P, 1 A)	(2, 3, 2, 3, 2)	(10857, 278812, 89497, 70989, 153316)	(4201, 17163, 6401, 2672, 2275)
6	michelelonghi	31	3	(1 A, 1 B, 1 P)	5	(1 A, 1 S, 1 B, 2 T)	(1, 1, 1, 1, 1)	(4172213, 4307717, 1145, 1716589, 539630)	(282, 425, 2158, 674, 4069)
7	msrvincent	22	3	(2 B, 1 T)	5	(4 P, 1 A)	(1, 1, 1, 1, 1)	(53288, 609125, 16624, 2618256, 10325)	(922, 1276, 4168, 91950, 1882)
8	miriamberkley	9	3	(1 B, 1 T, 1 H)	5	(1 P, 1 A, 2 B, 1 H)	(1, 1, 1, 1, 1)	(57289, 31965, 65574, 44249, 16979173)	(527, 293, 39857, 1385, 1048)
9	TheChuckMadel	12 6	3	(1 B, 1 P, 1 A)	5	(3 P, 1 T, 1 H)	(1, 1, 1, 1, 1)	(1209867, 723167, 539399, 598346, 1435742)	(2199, 318, 327, 96, 670)
10	RB_Recruitment	37	3	(1 B, 1 H, 1 T)	5	(2 B, 1 S, 1 A, 1 P)	(1, 1, 4, 1, 4)	(708, 1086384, 27807, 960, 3)	(1602, 3343, 168, 1274, 15)
11	OneWiretweets	14	3	(1 B, 1 T, 1 H)	5	(4 B, 1 S)	(1, 1, 1, 1, 1)	(2242, 2566271, 1714, 213, 231)	(2247, 371, 3739, 280, 680)
12	AdessoCreative	30	3	(1 B, 1 T, 1 S)	5	(4 B, 1 S)	(1, 5, 3, 2, 6)	(940207, 551725, 7671, 183933, 151651)	(12044, 27591, 3343, 48422, 26097)
...
30	jkaldor3	21	3	(1 S, 1 T, 1 B)	5	(3 S, 1 B, 1 T)	(1, 1, 1, 1, 1)	(131, 15491904, 17138, 61744, 43593)	(295, 1450, 1725, 480, 904)
31	Radiant_Body	31	3	(1 T, 1 P, 1 H)	5	(3 T, 2 P)	(1, 1, 1, 1, 1)	(148916, 25075, 324, 324, 35239)	(96352, 616, 212, 212, 2494)
32	RengsecondEng	11 0	3	(1 T, 1 P, 1 A)	5	(1 P, 2 B, 1 H, 1 T)	(1, 1, 1, 1, 1)	(2091177, 6461, 736767, 8010, 4563531)	(259, 2184, 464, 1666, 1999)
33	edyong209	10	3	(1 T, 1 B, 1 H)	5	(4 T, 1 A)	(1, 1, 1, 1, 1)	(4023039, 6054, 120, 2362, 160001)	(194, 5827, 1, 1975, 991)

No	1	2	3	4	5	6	7	8	9
34	mikeonitstuff	83	3	(1 T, 1 A, 1 P)	5	(3 P, 2 T)	(1, 1, 1, 1)	(35519, 49124, 21799, 575, 60)	(1695, 575, 996, 727, 63)
35	GeeksOnABeach	92	3	(1 T, 1 B, 1 P)	5	(3 B, 2 T)	(1, 1, 1, 1)	(297, 18378, 8597966, 1335, 42702)	(208, 2366, 893, 471, 2707)
36	kizel	93	3	(1 T, 1 P, 1 A)	5	(2 T, 2 P, 1 A)	(2, 2, 2, 1, 2)	(524878, 22536, 2085, 72906, 611435)	(2924, 4739, 2154, 7203, 4598)

Algoritma 1 Pengolahan data pengguna untuk diambil nilai *social trust*

(Input: idPerson

Proses: Melakukan operasi berkaitan dengan pengelolaan database berisi data pengguna dengan idPerson

Output: Hasil pengelolaan data pengguna dari database

Algoritma

CONSTRUCTOR

GETTER

SETTER

```
public void addFriend (person friend){
```

```
    add a friend element to List Friends
```

```
}
```

```
public void addTweet (tweet tweetMe){
```

```
    add a tweetMe element to List Tweets
```

```
}
```

```
public double computeL(){
```

```
    Declare pembagi
```

```
    Sum following and follower and add it into pembagi
```

```
    Add following divided by pembagi into L
```

```
    GET L
```

```
}
```

```
public double computeRT(person User){
```

```
    Declare pembagi
```

```
    Add retweet of User into pembagi
```

```
    Add retweet divided by pembagi into RT
```

```
    GET RT
```

```
}
```

```
public double computeTrust(){
```

```
    Multiply L with 0.5
```

```
    Sum it with RT
```

```
    Add it into trustValue
```

```
}
```

Proses di atas menghasilkan nilai kepercayaan yang disebut $trust_{u,f}$ antar pengguna dengan setiap teman. Kemudian nilai kepercayaan tersebut disimpan dalam database pada tabel relasi di mana setiap hubungan pengguna dengan temannya disimpan dalam tabel tersebut.

Algoritma 2 dan 3 merepresentasikan proses sentimen analisis konten *tweet* untuk diambil nilai *rating* pengguna terhadap topik berita tertentu.

Algoritma 2 Pengolahan konten tweet untuk menghasilkan rating

(Input: idTweet)

Proses: Melakukan operasi berkaitan dengan pengelolaan database berisi tweet dengan idTweet

Output: Hasil pengelolaan data tweet dari database

Algoritma

CONSTRUCTOR

GETTER

SETTER

Preprocessing with Lucene

```

public void calculateSImr(token tokenOfTweet){
    IF wcj of tokenOfTweet = 4 OR wcj of tokenOfTweet = 5 THEN
        Sum SImr[wcj of tokenOfTweet] with posScore of tokenOfTweet
    ELSE IF wcj of tokenOfTweet = 1 OR wcj of tokenOfTweet = 2 THEN
        Sum SImr[wcj of tokenOfTweet] with negScore of tokenOfTweet
    ELSE IF wcj of tokenOfTweet = 3 THEN
        Sum SImr[wcj of tokenOfTweet] with one
    }

public void calculateCountOfWcj(token tokenOfTweet){
    Initialize i to one
    FOR i=1 until i<=5 increment i
        IF wcj of tokenOfTweet = i THEN
            Sum countWcj[wcj of tokenOfTweet] with one
        ELSE
            Do nothing
    END OF FOR
    Sum nToken with one
}

public void calculateProbability(){
    Initialize i to one
    Initialize jum to nToken of current object
    FOR i=1 until i<=5 increment i
        Add countWcj[i] divided by jum to Pc[i]
    END OF FOR
}

public void calculateEmr(){
    Initialize i to one
    FOR i=1 until i<=5 increment i
        Add SImr[i] multiplied by Pc[i] to Emr[i]
        Sum sumOfEmr with Emr[i]
    END OF FOR
}

public void calculateEstimasi(){
    Initialize i to one
    FOR i=1 until i<=5 increment i
        Add Emr[i] divided by sumOfEmr to Emr2[i]
    END OF FOR
}

public double calculateRating(){
    Initialize i to one
    FOR i=1 until i<=5 increment i
        IF Emr[i] != zero THEN
            Sum SRmr with Emr2[i] multiplied by i
        ELSE
            Do nothing
    END OF FOR
}

```


Algoritma 3 Pengolahan tiap token yang menjadi bagian dari sebuah tweet

(Input: idToken

Proses: Melakukan operasi berkaitan dengan pengelolaan token dari sebuah tweet

Output: Hasil pengelolaan token-token dari sebuah tweet

Algoritma

Constructor

Getter

Setter

```
public void calculateWcj(){
    IF posScore is same as zero and negScore is same as zero THEN
        Add wcj to 3
    ELSE IF posScore is greater or same as zero and negScore is greater or same as
    zero THEN
        IF posScore is greater or same as negScore THEN
            IF posScore is greater or same as 0.5 THEN
                Add wcj to 5
            ELSE
                Add wcj to 4
        ELSE IF negScore is greater than posScore THEN
            IF negScore is greater or same as 0.5 THEN
                Add wcj to 1
            ELSE
                Add wcj to 2
        ELSE
            Do nothing
    }
```

Berikut contoh hasil pengolahan *implicit social trust* dan sentimen analisis *tweet* pada seorang pengguna yang dijadikan *input* untuk algoritma SVR:

Tabel 2 *Input* LIBSVM

SR	Min Trust	Max Trust	Min SR	Max SR
3,003436426116	0,479528259902	0,715147143462	3,01789264413	3,01789264413

Kolom SR menunjukkan nilai rating yang diberikan oleh pengguna tanpa terpengaruh dari preferensi teman. Rating tersebut didapatkan karena pengguna telah melakukan tweet yang berisi opini pengguna akan suatu artikel berita untuk topik tertentu. Min Trust merupakan nilai kepercayaan terkecil pada jejaring pertemanan pengguna yang melakukan rating terhadap topik berita yang sama. Max Trust merupakan nilai kepercayaan terbesar pada jejaring pertemanan pengguna yang melakukan rating terhadap topik berita yang sama. Min SR merupakan nilai rating terkecil yang diberikan teman pengguna akan suatu topik berita. Nilai ini didapatkan dari pengolahan tweet teman yang berisi opini teman akan suatu artikel berita tertentu. Max SR merupakan nilai rating terbesar yang diberikan teman pengguna akan suatu topik berita. Data tersebut kemudian diolah dengan memanfaatkan LIBSVM sehingga menghasilkan model untuk estimasi rating selanjutnya. Pada kasus ini digunakan epsilon bernilai 0,01.

Data uji diekstrak *implicit social trust*-nya dengan cara yang sama pada data latih. Hasil dari estimasi rating yang diberikan model dan rating sesungguhnya dari pengguna menjadi input pada Mean Absolute Percentage Error (MAPE). Tabel berikut menunjukkan hasil dari rating sebenarnya dengan estimasi rating dari sistem:

Tabel 3 Perbandingan Rating Hasil Sentimen Analisis Dan Sistem

Rating Sebenarnya	Rating dari Sistem
2,997402597	3,007040486
2,965517241	3,009829482
3,142857143	3,009768215
2,985407725	2,980500959
2,994108984	3,009289211
3,004115226	3,016571535
3,032067511	3,014307114
3,001733102	3,00913317
3,021479714	3,008090096
3,023575639	3,006003147
3,033962264	3,002943844
3	3,008652627

Hasil perhitungan dengan MAPE menunjukkan bahwa sistem memiliki tingkat akurasi yang tinggi di mana MAPE bernilai 0,8243075902233644%.

4. KESIMPULAN

Berdasarkan data yang sudah diolah dan penelitian yang telah dilakukan berkaitan dengan penelitian yang disusun, dapat diambil kesimpulan sebagai berikut:

1. Sistem Rekomendasi Berita dapat merekomendasikan berita dengan topik *art* untuk pengguna setelah dilakukan pengolahan *implicit social trust* pada pengguna dengan *username* "leoboudo" serta memanfaatkan algoritma *Support Vector Regression*.
2. Setelah dilakukan proses testing pada keenam pengguna utama didapatkan nilai *Mean Absolute Percentage Error* 0,8243075902233644%. Di mana hasil perhitungan MAPE ini menunjukkan bahwa sistem rekomendasi berita dapat menghasilkan rekomendasi berita yang sesuai karena semakin besar nilai MAPE yaitu mendekati nilai 100% maka semakin tidak akurat hasil yang diberikan sistem, sedangkan apabila nilai MAPE semakin kecil maka hasil yang diberikan sistem semakin akurat.

5. SARAN

Data yang digunakan untuk testing dalam penelitian ini tidak secara *realtime* sehingga diharapkan peneliti selanjutnya dapat mengambil data secara *realtime* dan dapat dikoneksikan langsung dengan akun Twitter pengguna. Penelitian selanjutnya juga dapat membuat translator untuk tweet sehingga dapat diaplikasikan pada bahasa Indonesia dan dapat diterapkan untuk berita dari Indonesia.

DAFTAR PUSTAKA

- [1] H. S. Husin, "News Recommendation Based On Web Usage And Web Content Mining," *29th Int. Conf. Data Eng. Work.*, pp. 326–329, 2013.
- [2] N. Jonnalagedda and S. Gauch, "Personalized News Recommendation Using Twitter," in *Proceedings - 2013 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Workshops, WI-IATW 2013*, 2013, pp. 21–25.
- [3] CNN, "ABOUT CNN.COM - CNN.com." [Online]. Available: <http://edition.cnn.com/about>.
- [4] D. H. Alahmadi and X.-J. Zeng, "ISTS: Implicit social trust and sentiment based approach to recommender systems," *Expert Systems with Applications*, vol. 42, no. 22, pp. 8840–8849, 2015.
- [5] C. Chaochao, Z. Jing, Z. Xiaolin, and C. Deren, "Recommender System Based on Social Trust Relationships," *E-bus. Eng. (ICEBE), 2013 IEEE 10th Int. Conf.*, pp. 32–37, 2013.
- [6] C. D. Manning, *Introduction to Information Retrieval*, vol. 1, 2015.
- [7] "SentiWordNet." [Online]. Available: <http://sentiwordnet.isti.cnr.it/>.
- [8] C. Chang and C. Lin, "LIBSVM: A Library for Support Vector Machines," pp. 1–39, 2013.
- [9] A.J. Smola and B. Schölkopf, "A Tutorial On Support Vector Regression," *Stat. Comput.*, vol. 14, pp. 199–222, 2004.
- [10] P. Swamidass, *Encyclopedia Of Production And Manufacturing Management*, Springer, 2000.