

BAB I

PENDAHULUAN

1.1 Latar Belakang

Berdasarkan data Kementerian komunikasi Dan Informatika(KOMINFO) pada tahun 2014 jumlah pengguna internet di Indonesia mencapai 88,3 juta orang. Angka tersebut menempatkan negara Indonesia pada posisi ke-6 pengguna internet di seluruh dunia. Hal tersebut menunjukkan masyarakat telah menggunakan internet di kehidupan sehari-hari. Internet menawarkan kemudahan akses dalam mendapatkan informasi, salah satunya pada website berita di internet. Konsumsi akses berita telah mengalami perubahan dari model tradisional seperti koran, majalah dan media cetak lainnya menjadi akses ribuan berita dari berbagai sumber di internet.

Top 25 Countries, Ranked by Internet Users, 2013-2018 millions						
	2013	2014	2015	2016	2017	2018
1. China*	620.7	643.6	669.8	700.1	736.2	777.0
2. US**	246.0	252.9	259.3	264.9	269.7	274.1
3. India	167.2	215.6	252.3	283.8	313.8	346.3
4. Brazil	99.2	107.7	113.7	119.8	123.3	125.9
5. Japan	100.0	102.1	103.6	104.5	105.0	105.4
6. Indonesia	72.8	83.7	93.4	102.8	112.6	123.0
7. Russia	77.5	82.9	87.3	91.4	94.3	96.6
8. Germany	59.5	61.6	62.2	62.5	62.7	62.7
9. Mexico	53.1	59.4	65.1	70.7	75.7	80.4
10. Nigeria	51.8	57.7	63.2	69.1	76.2	84.3
11. UK**	48.8	50.1	51.3	52.4	53.4	54.3
12. France	48.8	49.7	50.5	51.2	51.9	52.5
13. Philippines	42.3	48.0	53.7	59.1	64.5	69.3
14. Turkey	36.6	41.0	44.7	47.7	50.7	53.5
15. Vietnam	36.6	40.5	44.4	48.2	52.1	55.8
16. South Korea	40.1	40.4	40.6	40.7	40.9	41.0
17. Egypt	34.1	36.0	38.3	40.9	43.9	47.4
18. Italy	34.5	35.8	36.2	37.2	37.5	37.7
19. Spain	30.5	31.6	32.3	33.0	33.5	33.9
20. Canada	27.7	28.3	28.8	29.4	29.9	30.4
21. Argentina	25.0	27.1	29.0	29.8	30.5	31.1
22. Colombia	24.2	26.5	28.6	29.4	30.5	31.3
23. Thailand	22.7	24.3	26.0	27.6	29.1	30.6
24. Poland	22.6	22.9	23.3	23.7	24.0	24.3
25. South Africa	20.1	22.7	25.0	27.2	29.2	30.9
Worldwide***	2,692.9	2,892.7	3,072.6	3,246.3	3,419.9	3,600.2

Note: Individuals of any age who use the internet from any location via any device at least once per month; *excludes Hong Kong; **forecast from Aug 2014; ***includes countries not listed
Source: eMarketer, Nov 2014

Gambar 1.1 statistik pengguna internet di dunia (sumber gambar: Kominfo)

Banyaknya pengakses pada website berita dapat dimanfaatkan oleh pengelola dan pengembang sebuah situs berita untuk menarik pengunjung dan melakukan akses lebih lama. Karena berdasarkan statistik, website yang sering

dikunjungi di Indonesia berdasarkan situs Alexa adalah mesin pencari seperti Google atau Yahoo dan juga sosial media seperti Facebook. Sedangkan situs berita adalah Detik.com yang menempati urutan 9.

Nama	Peringkat Indonesia	Peringkat Global
Google.co.id	1	76
Google.com	2	1
Facebook.com	3	2
Youtube.com	4	3
Blogspot.com	5	24
Yahoo.com	6	4
Kaskus.co.id	7	291
WordPress.com	8	37
Detik.com	9	432
Ask.com	10	35

Gambar 1.2 Situs Web yang sering dikunjungi di Indonesia.

(sumber gambar: JurnalWeb)

Salah satu cara untuk membuat akses pengunjung lebih lama dari sebuah website adalah membuat sistem rekomendasi. Beberapa tahun terakhir sistem rekomendasi telah menjadi lazim digunakan untuk mengelola masalah data dan informasi (Wang et al. 2014). Sistem rekomendasi dapat memberikan saran atau pilihan artikel yang relevan untuk pengguna dalam mencari artikel berita pada sebuah website berita. Cabang ilmu pengetahuan yang dapat digunakan dalam membuat sistem rekomendasi salah satunya adalah data mining. Karena data mining memiliki potensi besar dalam penggalian ilmu yang bermanfaat dari sejumlah besar data sementara untuk pengambilan keputusan yang dinamis (Ltifi et al. 2016).

Berdasarkan penelitian Covington (2016), sistem rekomendasi menggunakan arsitektur Jaringan Syaraf Tiruan (JST) dapat digunakan dalam merekomendasikan video youtube dengan cara membagi dua bagian yaitu kandidat video dan daftar ranking dari video yang sudah ada secara efektif (Covington et al.

2016). Sedangkan menurut Yazdanfar (2013), sistem rekomendasi menggunakan *collaborative filtering* dapat memberikan saran URL kepada pengguna twitter berdasarkan kesamaan *hashtag* dan pengguna secara relevan dengan nilai Root Mean Squence Error(RMSE) sebesar 0,83 (Yazdanfar & Thomo 2013).

Salah satu algoritma lain dari data mining yang dapat digunakan dalam membangun sistem rekomendasi adalah *K-Nearest Neighbor*. Algoritma K-Neaerst Neighbor diterapkan dalam cabang ilmu pengetahuan lain seperti pengenalan pola, prediksi/rekomendasi, klasifikasi dan lain-lain. Penerapan lain dari metode K-Nearest Neighbor(KNN) digunakan secara online dan realtime untuk mengeksploitasi pada web-usage untuk mengidentifikasi aktivitas yang dilakukan oleh pengunjung atau klien yang nantinya akan dilakukan pencocokan kepada pengguna tertentu yang nantinya menjadi penyesuaian kebutuhan pengguna lain pada waktu-waktu tertentu (Adeniyi Q2 et al. 2014).

Teknik sistem rekomendasi yang dapat digunakan *Content Base Filtering(CBF)* dan memadukannya dengan metode KNN. Teknik CBF menentukan rekomendasi berdasarkan evaluasi dari pilihan pengguna di masa lampau yang kemudian diolah untuk mendapatkan rekomendasi untuk pengguna. Dengan menggunakan Teknik CBF (Isinkaye et al. 2015), data pengguna tidak terlalu berpengaruh dengan penentuan rekomendasi. Sehingga pada perancangan tidak membutuhkan user profile, walaupun sistem rekomendasi yang dihasilkan bersifat short time periode, yaitu sistem tidak dapat menentukan *interest* yang lebih mendalam dari pengguna.

Berdasarkan data dan uraian yang telah diuraikan dan penerapan metode KNN, maka penulis mencoba membuat sistem rekomendasi penentu artikel pada sebuah website yang relevan untuk pengguna atau pengunjung untuk memperoleh informasi. Sasaran dari sistem rekomendasi ini juga pemilik website tersebut, agar *traffic* website dan statistik pengunjung dapat naik sehingga dapat keuntungan dari mengelola website tersebut. Sistem yang akan dibuat dalam bentuk tugas akhir dengan judul “SISTEM REKOMENDASI PENENTU ARTIKEL PADA WEBSITE DENGAN MENGGUNAKAN METODE K-NEAREST NEIGHBOR.

1.2 Rumusan Masalah

Berdasarkan Latar belakang yang telah disusun, permasalahan yang akan diselesaikan adalah membangun sebuah sistem rekomendasi menggunakan metode K-Nearest Neighbor untuk artikel pada website yang relevan dengan artikel yang telah diakses pengguna dan mengukur keakuratan rekomendasi yang dihasilkan dalam bentuk Root Mean Squance Error (RMSE).

1.3 Tujuan

Tujuan dari pembuatan tugas akhir ini adalah sebagai berikut :

- a. Membangun plugin website yang akan digunakan sebagai media penyedia artikel dalam berbagai kategori
- b. Menerapkan metode *K-Nearest Neighbor* pada website yang telah dibangun untuk mendapatkan keakuratan dari rekomendasi dari artikel yang telah diakses.

1.4 Manfaat

Manfaat yang diinginkan dicapai dari pembuatan tugas akhir ini adalah :

- a. Dapat mengimplementasikan teknik dari data mining yaitu *K-Nearest Neighbor* untuk membuat sistem rekomendasi.
- b. Sebagai penyedia website dapat memperoleh trafik pengunjung yang tinggi karena pengunjung dapat mengakses website lebih lama
- c. Sebagai pengguna website mendapat kemudahan memilih akses berita lain dari rekomendasi tersebut tanpa harus mencari berita-berita lain.

1.5 Batasan Masalah

Berdasarkan latar belakang yang telah disusun agar penyelesaian tugas akhir ini tidak menyimpang maka penulis menentukan batasan masalah sebagai berikut

- a. Sistem rekomendasi untuk artikel adalah berdasarkan text mining dari judul, kategori artikel dari pengguna, yang dihasilkan dari metode *K-Nearest Neighbor* dalam bentuk Tag dari artikel.

- b. Rekomendasi yang dihasilkan pada pembuatan tugas akhir ini tidak berdasarkan pada user profile, tetapi berdasarkan tagline berita yang sedang diakses