

Prediksi Pengaruh Media Sosial terhadap Pilihan Mahasiswa pada Pilkada Serentak Menggunakan Data Mining

Amar P. Natasuwarna

STMIK Pontianak
Jurusan Sistem Informasi
e-mail: a.pegirosa@gmail.com

Abstrak

Media sosial telah memudahkan suatu kelompok manusia untuk saling mempengaruhi dan bertukar informasi. Apalagi sebagian masyarakat Indonesia terutama para mahasiswa sudah memiliki akun media sosial. Hal ini membuat para calon kepala daerah memanfaatkan media sosial ini sebagai sarana sosialisasi dan memperkenalkan program-programnya. Namun, selain mempunyai banyak sisi positif, media sosial juga dapat memberi dampak negatif dengan adanya keterbukaan informasi. Sehingga para calon kepala daerah harus dapat memanfaatkan media sosial untuk mempengaruhi calon pemilih termasuk para mahasiswa supaya dapat memilihnya pada saat pencoblosan. Tujuan penelitian ini adalah untuk memprediksi pengaruh media sosial terhadap pilihan mahasiswa pada pemilihan kepala daerah serentak. Adapun media sosial yang termasuk dalam penelitian ini adalah media sosial yang paling banyak dipergunakan oleh pengguna Internet yakni Facebook, Instagram, dan Twitter. Tahapan penelitian ini adalah menyebarkan kuisioner, melakukan preprocessing, membuat modeling algoritma ID3 dan C4.5 dengan perangkat lunak RapidMiner versi 8.1, analisis data dan menguji ketepatannya kedua algoritma tersebut. Hasil akurasi yang diperoleh dari pengolahan data mining dengan algoritma ID3 adalah 75%, sedangkan algoritma C4.5 adalah 70%.

Kata kunci: Media Sosial, Algoritma ID3, Algoritma C4.5, Pilkada

1. Pendahuluan

Tahun 2018 adalah tahun politik untuk 171 daerah di Indonesia [1]. Pemilihan kepala daerah atau yang disingkat dengan Pilkada adalah pemilihan walikota, bupati dan gubernur beserta wakilnya masing-masing serempak dilakukan pada tahun tersebut. Sebelum pelaksanaan Pilkada serentak, para calon kepala daerah melakukan sosialisasi dan mengirimkan berita-berita mengenai kelebihan mereka masing-masing melalui berbagai media termasuk media sosial yang merupakan media yang belum lama lahir dibandingkan dengan media lain yang sudah mapan seperti pemasangan

spanduk, billboard, surat kabar, radio, dan televisi. Media sosial sendiri diartikan sebagai forum online pertukaran informasi atau berita diantara kelompok masyarakat seperti kelompok keluarga atau kelompok pertemanan, selain itu media sosial mampu meningkatkan pengetahuan masyarakat dan memberi pengaruh dalam partisipasi politik sehubungan adanya saling percaya pada kelompok tersebut [2]. Sebagai teknologi baru, media sosial selalu diasosiasikan dengan kalangan muda termasuk mahasiswa, terbukti 89% pemuda memiliki akun media sosial, hal ini dilain pihak dapat dimanfaatkan oleh para calon kepala daerah untuk memobilisasi informasi dengan mengajak kalangan muda untuk aktif dalam partisipasi politik dengan cara memilihnya pada saat pencoblosan [3]. Pada partisipasi politik di Indonesia tahun 2014, sebanyak 30% pemilih adalah pemuda di bawah 30 tahun, sehingga membuat para calon kepala daerah semakin mengarahkan komunikasi politiknya melalui saluran media sosial, walaupun pada sisi yang lain, media sosial tidak hanya memberitakan hal positif tetapi juga hal-hal yang negatif dari perilaku para calon kepala daerah seperti sejarah pernah melakukan korupsi dan kolusi, dan adanya kampanye hitam dapat juga mempengaruhi pilihan para pemuda [4]. Untuk memprediksi apakah media sosial dapat mempengaruhi para pemuda atau mahasiswa dalam memilih kepala daerah, salah satunya adalah menggunakan metoda data mining.

Data mining merupakan teknik yang paling sering digunakan untuk menemukan pola-pola yang tidak diketahui sebelumnya, salah satunya adalah melakukan prediksi dengan cara data-data tersebut diketahui terlebih dahulu kelas masing-masing dari sampel sehingga menghasilkan model *classifier* yang dapat melakukan prediksi ketika diberikan sampel baru [5]. Dalam proses melakukan prediksi, data mining klasifikasi diantaranya menggunakan algoritma ID3 dan C4.5. Kedua algoritma ini menghasilkan pohon keputusan. Kedua algoritma ini sangat populer karena mempunyai keunggulan diantaranya adalah mempunyai akurasi yang baik dan hasil visualisasi pohon keputusan sangat mudah untuk dianalisis [6]. Pohon keputusan menghasilkan informasi yang dapat mendukung penggunaannya dalam membantu pengambilan keputusan. Hasil visualisasi yang ditampilkan dalam bentuk rangkaian *leaf node* yang menghasilkan *rule*. Perhitungan algoritma ID3

menggunakan persamaan *Entropy* dan *Information Gain* Sedangkan algoritma C4.5 adalah pengembangan dari ID3 yang diperkenalkan oleh Ross Quinlan, dengan pertama-tama melakukan perhitungan yang sama dengan memperoleh *Entropy*, selanjutnya dilakukan perhitungan *Gain Ratio* dengan terlebih dahulu mengetahui *Split Information*. Kelebihan dari C4.5 terhadap ID3 adalah dapat mengatasi *missing data* dan data kontinu [6][7].

Sudah banyak penelitian mempergunakan data mining yang berkaitan dengan aktivitas mahasiswa diantaranya melakukan prediksi terhadap performa akhir mahasiswa [8], prediksi mahasiswa dropout [9], prediksi pesan persuasif mempengaruhi perilaku mahasiswa [10]. Namun, penelitian mengenai hubungan antara mahasiswa dan media sosial yang dapat mempengaruhi pilihan mahasiswa dengan menggunakan data mining belum pernah ditemukan. Tujuan penelitian ini adalah melakukan prediksi pengaruh media sosial terhadap pilihan mahasiswa STMIK Pontianak pada Pilkada serentak yang diadakan pada bulan Juni 2018 dengan mempergunakan data mining.

Lokasi penelitian adalah STMIK Pontianak. Data diperoleh secara random berjumlah 100 dari berbagai angkatan mahasiswa yang belum lulus sarjana. Data diambil dari hasil survey dengan memberikan kuisioner untuk memperoleh data personal mahasiswa, media-media sosial yang dipergunakan, dan data mengenai apakah ada pengaruh media sosial terhadap pemilihan kepala daerah kepada mahasiswa STMIK Pontianak. Adapun data personal mahasiswa yang pernah dipergunakan pada penelitian data mining terdahulu, dan data personal ini juga dipergunakan pada penelitian ini adalah berupa atribut jenis kelamin dan usia [10]. Sementara itu, dari sekian banyak media sosial yang ada, dipergunakan sebagai bagian dari atribut penelitian adalah media sosial yang paling banyak dipergunakan oleh pengguna Internet yakni Facebook, Instagram, dan Twitter [11][12].

Sebagaimana proses prediksi pada data mining harus mempunyai atribut sebagai label atau target yang kemudian dikelompokkan menjadi beberapa kelas. Sebagai atribut label atau target adalah Pengaruh Medsos yang menyatakan ada pengaruh atau tidak media sosial terhadap pilihan mahasiswa. Untuk melakukan prediksi, metoda data mining yang dipergunakan adalah data mining klasifikasi dengan algoritma ID3 dan C4.5.

Tahapan penelitian ini adalah: pertama, menyebarkan kuisioner untuk memperoleh data yang diperlukan. Kedua melakukan *preprocessing* untuk memastikan bahwa data yang diperoleh siap untuk dilakukan pengolahan data. Ketiga, membagi data menjadi data training dan data testing. Keempat, membuat modeling dari data training dengan menggunakan perangkat lunak RapidMiner versi 8.1. Pada langkah ini, perangkat lunak RapidMiner melakukan perhitungan *Entropy* dan *Information Gain* untuk ID3 dan melakukan perhitungan *Entropy* dan *Gain Index* untuk C4.5. Langkah kelima melakukan analisis data dari hasil keluaran dalam bentuk visualisasi pohon keputusan. Terakhir, model pohon

keputusan tersebut diuji ketepatannya dengan memasukkan data testing.

Tahapan algoritma ID3 sebagai berikut [13], pertama menghitung *Entropy* dan *Information Gain*. Kedua, atribut dengan *Information Gain* terbesar menjadi menjadi simpul. Ketiga, menghitung kembali nilai *Entropy* dan *Information Gain* tiap-tiap atribut dengan tidak mengikutsertakan atribut yang terpilih menjadi simpul di tahap sebelumnya. Proses berakhir apabila semua data telah mempunyai kelas.

Tahapan algoritma C4.5 sebagai berikut [14], pertama menghitung *Entropy* dan *Gain Index*. Kedua, atribut dengan *Gain Index* terbesar menjadi menjadi simpul. Ketiga, menghitung kembali nilai *Entropy* dan *Gain Index* tiap-tiap atribut dengan tidak mengikutsertakan atribut yang terpilih menjadi simpul di tahap sebelumnya. Proses berakhir apabila semua data telah mempunyai kelas.

Adapun persamaan *Entropy* dan *Information Gain*, sebagai berikut [9][13][16].

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \dots\dots\dots (1)$$

Dimana S: himpunan kasus, A: atribut, V: menyatakan suatu nilai yang mungkin untuk atribut A, |Si|: jumlah kasus pada partisi ke-I, |S|: jumlah kasus dalam S, (Si): *entropy* untuk sampel-sampel yang memiliki nilai i.

Untuk mendapatkan *Information Gain* perlu dihitung *Entropy* sebagai mana persamaan (2) berikut.

$$Entropy(S) = \sum_{i=1}^n - pi * \log_2 pi \dots\dots\dots (2)$$

Dimana S: himpunan kasus, n: jumlah partisi S, pi: proporsi dari Si terhadap S.

Selanjutnya untuk C4.5 menghitung *Gain Ratio* sebagai mana persamaan (3). Namun untuk menghitung *Gain Ratio* diperlukan *Split Information* sebagaimana persamaan (4)[11].

$$Gain Ratio(S, X) = \frac{Information Gain(X)}{Split Information(S, X)} \dots\dots\dots (3)$$

$$Split Information(S, X) = - \sum_{i=1}^y \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|} \dots\dots\dots (4)$$

Namun sebelum proses pengolahan data dengan algoritma ID3 dan C4.5, dilakukan terlebih dahulu enam tahap yakni tahap pengumpulan data, tahap pembersihan data, tahap integrasi dan seleksi data, tahap transformasi data, terakhir tahap mining yang menghasilkan pengetahuan. Tahap pembersihan data hingga tahap transformasi disebut dengan tahap *preprocessing* [11].

Tahap *preprocessing* dilakukan apabila data yang diperoleh ternyata mengandung data kosong, selain itu perlu dilakukan pengurangan, mengandung noise, format yang tidak konsisten, dan memerlukan transformasi dari data kontinyu menjadi data diskrit. Sampel yang mengandung data kosong dan *noise* tidak dipergunakan untuk proses pengolahan data.

Selanjutnya, tahapan yang paling menentukan untuk data mining adalah proses mining berupa modeling menggunakan perangkat lunak RapidMiner versi 8.1. Terakhir, hasil dari data training yang telah menghasilkan model diuji dengan data testing. Berikut adalah persamaan pengujian yang menghasilkan akurasi sebagai mana persamaan(5) [17].

$$Akurasi (%) = \frac{Jumlah\ data\ benar}{Jumlah\ data\ yang\ diuji} \times 100\% \dots\dots (5)$$

2. Pembahasan

Data penelitian diambil dengan jumlah sampel yang diperoleh secara random dari 100 mahasiswa dari semua angkatan STMIK Pontianak terdiri-dari 8 atribut termasuk Jenis Kelamin, Usia, Facebook, Instagram, Twitter, dan Pengaruh Medsos sebagaimana yang ditunjukkan pada Tabel 1 berikut.

Tabel 1. Diskripsi Atribut

Nama Atribut	Deskripsi	Kemungkinan Nilai	Jumlah Pilihan
Jenis Kelamin	Jenis kelamin mahasiswa	(Pria, Wanita)	(73,27)
Usia	Kategori usia mahasiswa	(<20, >=20)	(45,55)
Facebook	Menggunakan media Facebook	(Ya, Tidak)	(81,18)
Instagram	Menggunakan media Instagram	(Ya, Tidak)	(94,6)
Twitter	Menggunakan media Twitter	(Ya, Tidak)	(24,76)
Pengaruh Medsos	Pengaruh media sosial terhadap pilihan mahasiswa	(Ya, Tidak)	(52,48)

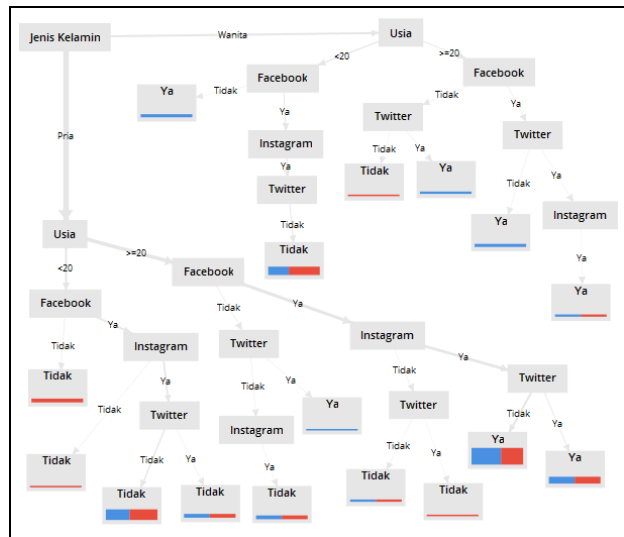
Berikut adalah penjelasan mengenai masing-masing atribut pada Tabel 1. Jenis kelamin adalah jenis kelamin mahasiswa sebagai informasi data personal. Kemungkinan nilai pilihan Pria dan Wanita. Usia adalah usia mahasiswa yang dibagi atas dua kelompok yakni kelompok dengan simbol <20 adalah usia dari 17 tahun hingga 19 tahun dan kelompok dengan simbol >=20 adalah usia mulai 20 tahun hingga 23 tahun. Atribut Tipe Medsos adalah jenis media sosial yang dipergunakan oleh mahasiswa. Untuk pertanyaan ini, mahasiswa diberi beberapa pertanyaan mengenai Tipe Medsos yang terdiri dari Facebook, Instagram, dan Twitter. Kemungkinan nilainya masing-masing adalah Ya dan Tidak. Pengaruh Medsos – adalah atribut label atau target yang menyatakan apakah ada pengaruhnya berbagai berita dan informasi yang muncul pada media sosial terhadap pilihan mahasiswa pada saat pelaksanaan pilkada serentak 2018. Kemungkinan nilainya adalah Ya dan Tidak.

Tabel 2. Contoh Himpunan Data

Jenis Kelamin	Usia	Facebook	Instagram	Twitter	Efektif
Pria	>=20	Ya	Tidak	Tidak	Ya
Pria	>=20	Tidak	Ya	Tidak	Tidak
Wanita	>=20	Tidak	Tidak	Ya	Tidak
Pria	<20	Tidak	Tidak	Ya	Ya
Wanita	>=20	Ya	Ya	Tidak	Tidak

Wanita	<20	Ya	Tidak	Ya	Ya
Wanita	<20	Tidak	Tidak	Ya	Ya

Tabel 2 menunjukkan contoh dari hasil preprocessing setelah data diperoleh dari hasil survey. Preprocessing yang dilakukan pertama adalah memeriksa apakah terdapat data kosong atau noise. Apabila ditemukan data kosong atau noise sampel tersebut tidak diikuti pada proses selanjutnya. Tahap berikut preprocessing yang dilakukan adalah melakukan transformasi data pada atribut usia dari data kontinyu menjadi data diskrit yang dapat diolah dengan menggunakan kedua algoritma tersebut, diperoleh dua kelompok. Dengan selesainya preprocessing, data selanjutnya dilakukan proses mining menggunakan perangkat lunak RapidMiner versi 8.1. Dari 100 sampel yang diperoleh secara random, dibagi atas dua bagian. Pertama 80 sampel digunakan sebagai data training, kemudian 20 sampel digunakan sebagai data testing. Pertama-tama proses data mining menggunakan algoritma ID3 dengan menggunakan perangkat lunak RapidMiner versi 8.1.



Gambar 1. Hasil Proses Algoritma ID3

Visualisasi pohon keputusan seperti yang terlihat pada Gambar 1 di atas merupakan hasil pengolahan data mining algoritma ID3 dengan menggunakan perangkat lunak RapidMiner versi 8.1, visualisasi tersebut dapat juga dinyatakan ke dalam deskripsi ID3 Tree yang dapat memperlihatkan jumlah kelas dari masing-masing atribut sebagai berikut.

ID3 Tree

- Jenis Kelamin = Pria
 - | Usia = <20
 - | | Facebook = Tidak: Tidak {Ya=0, Tidak=4}
 - | | Facebook = Ya
 - | | | Instagram = Tidak: Tidak {Ya=0, Tidak=1}
 - | | | Instagram = Ya
 - | | | | Twitter = Tidak: Tidak {Ya=6, Tidak=7}
 - | | | | Twitter = Ya: Tidak {Ya=2, Tidak=2}
 - | | Usia = >=20
 - | | Facebook = Tidak

```

| | | Twitter = Tidak
| | | Instagram = Ya: Tidak {Ya=2, Tidak=2}
| | | Twitter = Ya: Ya {Ya=1, Tidak=0}
| | Facebook = Ya
| | Instagram = Tidak
| | | Twitter = Tidak: Tidak {Ya=1, Tidak=1}
| | | Twitter = Ya: Tidak {Ya=0, Tidak=1}
| | Instagram = Ya
| | | Twitter = Tidak: Ya {Ya=12, Tidak=9}
| | | Twitter = Ya: Ya {Ya=4, Tidak=4}
Jenis Kelamin = Wanita
| Usia = <20
| | Facebook = Tidak: Ya {Ya=3, Tidak=0}
| | Facebook = Ya
| | | Instagram = Ya
| | | Twitter = Tidak: Tidak {Ya=4, Tidak=6}
| | Usia = >=20
| | Facebook = Tidak
| | | Twitter = Tidak: Tidak {Ya=0, Tidak=1}
| | | Twitter = Ya: Ya {Ya=2, Tidak=0}
| | Facebook = Ya
| | | Twitter = Tidak: Ya {Ya=3, Tidak=0}
| | | Twitter = Ya
| | | Instagram = Ya: Ya {Ya=1, Tidak=1}
    
```

Berdasarkan Gambar 1 dan ID3 Tree di atas dengan output Pengaruh Medsos dengan klasifikasi Ya atau Tidak, diperoleh *root node* adalah Jenis Kelamin, dengan jumlah *leaf nodes* atau *rule* adalah sejumlah 16.

Dari analisa pohon keputusan pada Gambar 1, pada atribut jenis kelamin pria dan usia kelompok <20 tahun (dominan Tidak): 4 *rule* semua menyatakan kelas Tidak. Selanjutnya pada atribut Jenis Kelamin Pria dan atribut kelompok usia =>20 tahun (tidak ada yang dominan): 6 *rule*, dimana 3 *rule* menyatakan kelas Ya, 3 *rule* menyatakan kelas Tidak. Pada atribut Jenis Kelamin Wanita dan kelompok usia <20 tahun (tidak ada yang dominan): 2 *rule* dimana 1 *rule* menyatakan Ya, 1 *rule* menyatakan Tidak. Sedangkan pada atribut Jenis Kelamin Wanita dan atribut kelompok =>20 tahun (dominan Ya): 4 *rule*, 3 *rule* menyatakan Ya, 1 *rule* menyatakan Tidak.

Sehubungan dengan Tipe Media Sosial, pada pengguna Facebook (dominan Tidak): 10 *rule*, 6 *rule* menyatakan Tidak, 4 menyatakan Ya. Pada pengguna Instagram (dominan Tidak): 7 *rule*, 4 *rule* menyatakan Tidak, 3 *rule* menyatakan Ya. Sedangkan pada pengguna Twitter (dominan Ya): 7 *rule*, 4 *rule* menyatakan Ya, 3 *rule* menyatakan Tidak.

Dari pengujian ketepatan antara modeling hasil pengolahan data training dan sampel baru pada data testing, maka dari pengujian yang dilakukan oleh RapidMiner versi 8.1 diperoleh hasil akurasi seperti pada Tabel 3.

Tabel 3. Akurasi ID3

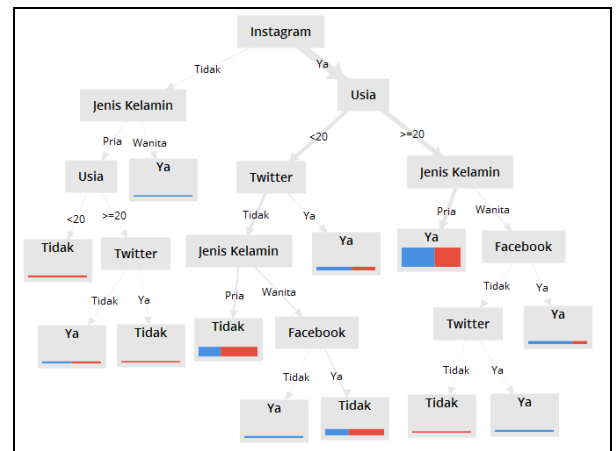
	True Ya	True Tidak	Class Precision
Prediction Ya	8	2	80%

Prediction	3	7	70%
Tidak			
Class	72,73%	77,78%	
Recall			

Berdasarkan Tabel 3, maka diperoleh Class Precision 80% dan 70%, Class Recall 72,73% dan 77,78%. Diperoleh Akurasi dengan menggunakan persamaan (5) untuk algoritma ID3 sebagai berikut:

$$\text{Akurasi} = \frac{8+7}{20} \times 100\% = 75\%$$

Selanjutnya proses data mining menggunakan algoritma C4.5.



Gambar 2. Hasil Proses Algoritma C4.5

Visualisasi pohon keputusan pada Gambar 2 sebagai hasil pengolahan data mining algoritma C4.5 dengan perangkat lunak RapidMiner versi 8.1 dapat dinyatakan ke dalam deskripsi C4.5 Tree yang dapat memperlihatkan jumlah kelas dari masing-masing atribut sebagai berikut.

C4.5 Tree

```

Instagram = Tidak
| Jenis Kelamin = Pria
| | Usia = <20: Tidak {Ya=0, Tidak=2}
| | Usia = >=20
| | | Twitter = Tidak: Ya {Ya=1, Tidak=1}
| | | Twitter = Ya: Tidak {Ya=0, Tidak=1}
| Jenis Kelamin = Wanita: Ya {Ya=1, Tidak=0}
Instagram = Ya
| Usia = <20
| | Twitter = Tidak
| | | Jenis Kelamin = Pria: Tidak {Ya=6, Tidak=10}
| | | Jenis Kelamin = Wanita
| | | | Facebook = Tidak: Ya {Ya=2, Tidak=0}
| | | | Facebook = Ya: Tidak {Ya=4, Tidak=6}
| | | Twitter = Ya: Ya {Ya=3, Tidak=2}
| | Usia = >=20
| | | Jenis Kelamin = Pria: Ya {Ya=19, Tidak=15}
| | | Jenis Kelamin = Wanita
| | | | Facebook = Tidak
| | | | | Twitter = Tidak: Tidak {Ya=0, Tidak=1}
| | | | | Twitter = Ya: Ya {Ya=2, Tidak=0}
| | | Facebook = Ya: Ya {Ya=3, Tidak=1}
    
```

Berdasarkan Gambar 5 dan C4.5 Tree di atas dengan output Pengaruh Medsos dengan klasifikasi Ya atau Tidak, diperoleh *root node* adalah Instagram, dengan jumlah *leaf nodes* atau *rule* adalah sejumlah 12. Dari analisa pohon keputusan pada Gambar 2, pada Tipe Media Sosial, pada pengguna Instagram (dominan Ya): 8 *rule*, 5 *rule* menyatakan Ya, 3 menyatakan Tidak. Pada pengguna Facebook (tidak ada yang dominan): 2 *rule*, 1 *rule* menyatakan Ya, 1 *rule* menyatakan Tidak. Sedangkan pada pengguna Twitter (dominan Ya): 3 *rule*, 2 *rule* menyatakan Ya, 1 *rule* menyatakan Tidak. Selanjutnya, pada atribut jenis kelamin pria dan usia kelompok <20 tahun (dominan Tidak): 2 *rule* semua menyatakan kelas Tidak. Pada atribut Jenis Kelamin Pria dan atribut kelompok usia =>20 tahun (dominan Ya): 3 *rule*, dimana 2 *rule* menyatakan kelas Ya, 1 *rule* menyatakan kelas Tidak. Pada atribut Jenis Kelamin Wanita dan kelompok usia <20 tahun (dominan Ya): 3 *rule* dimana 2 *rule* menyatakan Ya, 1 *rule* menyatakan Tidak. Sedangkan pada atribut Jenis Kelamin Wanita dan atribut kelompok usia =>20 tahun (dominan Ya): 4 *rule*, 3 *rule* menyatakan Ya, 1 *rule* menyatakan Tidak. Dari pengujian ketepatan antara modeling hasil pengolahan data training dan sampel baru pada data testing, maka dari pengujian yang dilakukan oleh RapidMiner versi 8.1 pada diperoleh hasil akurasi seperti pada Tabel 4.

Tabel 4. Akurasi C4.5

	True Ya	True Tidak	Class Precision
Prediction Ya	9	4	69,23%
Prediction Tidak	2	5	71,43%
Class Recall	81,82%	55,56%	

Berdasarkan Tabel 4, maka diperoleh Class Precision 69,23% dan 71,43%, Class Recall 81,82% dan 55,56,78%. Diperoleh Akurasi dengan menggunakan persamaan (5) untuk algoritma C4.5 adalah sebagai berikut:

$$\text{Akurasi} = \frac{9+5}{20} \times 100\% = 70\%$$

Tabel 5. Akurasi Klasifier

Algoritma	Correctly (%)	Incorrectly (%)
ID3	75%	25%
C4.5	70%	30%

Berdasarkan Tabel 5 menunjukkan akurasi klasifier pada pohon keputusan. Dimana akurasi algoritma ID3 adalah 75% dan akurasi C4.5 sebesar 70%.

3. Kesimpulan

Dari hasil analisis kedua pohon keputusan hasil proses algoritma ID3 dan C4.5 ditemukan kesamaan bahwa jenis kelamin pria yang berusia dari 17 hingga 19 tahun pilihan politik mereka untuk memilih calon kepala

daerah dominan tidak dipengaruhi dari media sosial, sebaliknya pada jenis kelamin wanita yang berusia dari 20 hingga 23 tahun pilihan politik mereka untuk memilih calon kepala daerah dominan dipengaruhi oleh media sosial. Sedangkan dari sisi tipe media sosial, pengguna Twitter pilihan politiknya mereka untuk memilih calon kepala daerah, dominan dipengaruhi oleh media sosial. Sehubungan dengan performa, proses algoritma ID3 dan C4.5 menunjukkan hasil yang cukup baik untuk memprediksi pengaruh media sosial pada mahasiswa untuk memilih calon kepala daerah. Dengan menggunakan data training dan data testing diperoleh akurasi menggunakan data mining klasifikasi algoritma ID3 adalah 75%, sedangkan C4.5 adalah 70%. Untuk kelanjutan, ada beberapa hal yang dapat dilakukan untuk penelitian berikutnya yakni dengan memperluas ruang lingkup penelitian. Penelitian tidak hanya dilakukan sebatas pada mahasiswa STMIK Pontianak saja tetapi dapat dilakukan pada beberapa perguruan tinggi yang ada di kota Pontianak. Selain itu, penelitian dapat juga dilakukan dengan menggunakan algoritma lain yang ada pada data mining klasifikasi.

Daftar Pustaka

- [1]. M.A. Aziz, "Pilkada Serentak melalui DPRD: Sebuah Gagasan Mewujudkan Pilkada Demokratis Perspektif Pancasila dan UUD 1945", Politik Indonesia: Indonesian Political Science Review, Vol. 1, No. 2, pp. 154-170, 2016.
- [2]. S. Boulianne, "Social Media Use and Participation: A Meta-analysis of Current Research", Information, Communication & Society, Vol. 18, No. 5, pp. 524-538, 2015.
- [3]. A. Salman, & S. Saad, "Online political participation: A study of youth usage of new media", Mediterranean Journal of Social Sciences, Vol. 6, No. 4, p. 88, 2015.
- [4]. A. Kholid, R. Husein, & D. Mutiarin, "The Influence of Social Media Towards Student Political Participation During the 2014 Indonesian Presidential Election", Journal of Government and Politics, Vol. 6, No. 2, pp. 246-264, 2015.
- [5]. A.B. Ahmed, & I.S. Elaraby, "Data Mining: A prediction for Student's Performance Using Classification Method", World Journal of Computer Application and Technology, Vol. 2, No. 2, pp. 43-47, Feb., 2014.
- [6]. M.F. Ami, "Penerapan Algoritma Cart Untuk Memprediksi Status Kelulusan Mahasiswa", Jutisi, Vol. 5, No. 3, pp. 1215-1222, Feb., 2017.
- [7]. N. Arista, M. Bettiza, & H. Kurniawan, "Penerapan Algoritma ID3 dan C4. 5 dalam Menemukan Hubungan Data Awal Masuk Mahasiswa dengan Prestasi Akademik", Universitas Maritim Raja Ali Aji Tanjungpinang, pp. 1-10, 2014.
- [8]. N.A. AL-Fakhry, "Summarizing data by using data mining techniques a comparative by using C4. 5 and C5. 0 algorithms", International Education and Research Journal, Vol.2, No.4 , pp. 8-12, 2016.
- [9]. D. Heredia, Y. Amaya, & E. Barrientos, "Student dropout predictive model using data mining techniques", IEEE Latin America Transactions, Vol. 13, No. 9, pp.3127-3134, 2015.
- [10]. Y. Promdee, S. Kasemvilas, N. Phangsuk, & R. Yodthasarn, "Predicting Persuasive Message for Changing Student's Attitude Using Data Mining", InPlatform Technology and Service (PlatCon), 2017 International Conference, pp. 1-5, Feb., 2017.
- [11]. F. Rohman, "Analisis Meningkatnya Kejahatan Cyberbullying dan Hatespeech Menggunakan Berbagai Media Sosial dan Metode Pencegahannya," Seminar Nasional Ilmu Pengetahuan dan Teknologi Komputer, Vol. 4, No. 1, pp. 383-388, Feb., 2018.
- [12]. A. N. Viani, B. Santoso, "Media Baru dan Partisipasi Politik (Pengaruh Twitter Terhadap Tingkat Partisipasi Politik Remaja dalam Pilkada Serentak 2015 pada Mahasiswa Fakultas Ilmu Komunikasi dan Informatika Universitas Muhammadiyah

Seminar Nasional Sistem Informasi dan Teknologi Informasi 2018

SENSITEK 2018

STMIK Pontianak, 12 Juli 2018

- Surakarta Angkatan 2014)", Universitas Muhammadiyah Surakarta, pp. 1-15, 2015.
- [13]. A. E. Amalia, & M. Z. Naf'an, "Implementasi Algoritma ID3 untuk Klasifikasi Performansi Mahasiswa (Studi Kasus ST3 Telkom Purwokerto)", *Semnasteknomedia Online*, Vol. 5, No.1, p. 2-1, 2017.
- [14]. B. Novianti, T. Rismawan, S. Bahri, "Implementasi Data Mining dengan Algoritma C4.5 untuk Penjurusan Siswa (Studi Kasus SMA Negeri 1 Pontianak)", *Jurnal Coding Sistem Komputer Universitas Tanjungpura*. Vol. 4, No.3, pp. 75-84, Oct., 2016.
- [15]. B. A. Pereira, P. Anusha, & C. Fernandes. "A Comparative Analysis of Decision Tree Algorithms for Predicting Student's Performance." *International Journal of Engineering Science*, Vol.7, No.4, pp. 10489-10492, 2017.
- [16]. M. R. Azkia, Mayadi, & M. Huda, Kusriani, "Sistem Pendukung Keputusan Pemilihan Program Studi Calon Peserta Didik Menggunakan Metode Pohon Keputusan C4.5 (Studi Kasus: Prodi Universitas Amikom Yogyakarta)", *Seminar Nasional Teknologi Informasi dan Multimedia*, pp. 2.8-43-2.8-48, Feb., 2018.
- [17]. T.D. Utama, S.W. Sihwi, A. Doewes, "Implementasi Algoritma Iterative Dichotomiser 3 pada Penyeleksian Program Mahasiswa Wirausaha UNS", *ITSmart: Jurnal Teknologi dan Informasi*, Vol. 3, No.2, pp. 74-82, Aug., 2016.