

ANALISA PENCAPAIAN HAFALAN AL-QUR'AN DALAM CAMPTAHFIZH DENGAN ALGORITMA C.45 DAN PENERAPANNYA

Muchlis^{1*}, Taufiqurochman²

^{*12}Pasca Sarjana Ilmu Komputer, STMIK Nusamandiri, Jakarta
Jalan Kramat Raya No.18 Kwitang, 10420

^{*}E-mail : muchlis.re@gmail.com.

ABSTRAK

Tujuan penelitian ini adalah untuk menganalisa dan mengukur pencapaian hafalan santri dalam menghafal Al-Qur'an pada kegiatan CampTahfizh dengan Algoritma C.45. Dari hasil analisa akan di dapatkan pohon keputusan yang akan diterapkan pada aplikasi berbasis web yang dapat di manfaatkan dalam memprediksi pencapaian hafalan al-Qur'an setiap kegiatan CampTahfizh.

Kata kunci: C.45, hafalan, Al-Qur'an

ABSTRACT

The purpose of this study is to analyze and measure the achievement of recitation of santri in memorizing Al-Qur'an on CampTahfizh activities with Algorithm C.45. From the analysis results will get a decision tree that will be applied to web-based applications that can be utilized in predicting the achievement of memorization of the Qur'an every CampTahfizh activity.

Keywords : C.45, memorizing, Al-Qur'an

PENDAHULUAN

Pendidikan di Indonesia sangat pesat perkembangannya, baik dari pendidikan umum, pendidikan khusus atau luar biasa. Pendidikan merupakan salah satu wadah yang bertujuan untuk membentuk karakter manusia secara utuh. Melalui pendidikan manusia dapat mengembangkan potensi-potensi diri agar mampu bersaing dan bermanfaat bagi dirinya, keluarga, masyarakat serta bangsa dan negaranya. Sebagaimana tertera di Undang-undang Dasar Negara Republik Indonesia di alinea ke empat (Undang-undang No 20 Tahun 2003).

Menurut UU No 20 Tahun 2003 pendidikan dibagi menjadi tiga, pertama pendidikan formal yaitu pendidikan yang terstruktur dan berjenjang yang terdiri atas pendidikan dasar, pendidikan menengah dan pendidikan tinggi. Kedua, non formal yaitu pendidikan diluar pendidikan formal yang dapat dilaksanakan secara terstruktur dan berjenjang, terakhir adalah pendidikan informal yaitu jalu pendidikan keluarga dan lingkungan yang berbentuk kegiatan secara mandiri (suprijanto:2005).

CampTahfizh adalah program yang dilaksanakan di dalam pendidikan formal,

yaitu pondok pesantren. Program ini menggukan model karantina seluruh santri dalam lingkungan yang tidak bersentuhan kegiatan pendidikan sekolah seperti biasanya. Peserta *CampTahfizh* hanya melakukan menghafal al-Qur'an tanpa mengikuti kegiatan umum santri lainnya.

Peserta CampTahfizh berada dalam karantina selama 3 bulan, terhitung dari awal datang ke pesantren. Setelah CampTahfizh selesai, peserta memperoleh hasil yang berbeda-beda.

Beberapa tahun ini, banyak universitas negeri maupun swasta menerapkan ketentuan dalam menerima mahasiswa baru, yaitu jalur khusus bagi calon mahasiswa yang memiliki hafalan al-Qur'an untuk menjadi mahasiswa di universitas tersebut. Untuk itu, perlu adanya langkah yang dilakukan oleh sekolah atau pesantren yang memang di kurikulumnya memiliki program *tahfizh Qur'an*.

Langkah-langkas tersebut dapat diterapkan pada metode pembelajaran maupun perumusan kurikulum pada sekolah tersebut yang menerapkan hafalan Qur'an.

Dalam dunia pendidikan, data yang berlimpah dan berkesinambungan mengenai siswa yang dibina dan alumni terus dihasilkan.

Menurut Jing (2004) dan Merceron (2005) dalam Ayub (2007), data yang berlimpah membuka peluang diterapkannya data mining untuk pengelolaan pendidikan yang lebih baik dan data mining dalam pelaksanaan pembelajaran berbantuan komputer yang lebih efektif.

Banyak artikel atau jurnal yang membahas tentang kelulusan siswa atau tepat waktunya siswa dalam menempuh pendidikan di sekolah. Banyak juga artikel yang meneliti tentang ketepatan waktu lulus kuliah menggunakan algoritma C4.5.

Didalam artikel ini kami akan menganalisa hasil dari program yang diterapkan salah satu sekolah nasional yang menerapkan program hafalan al-Qur'an. Kami akan menganalisa hasil dari pencapaian hafalan al-Qur'an menggunakan metode algoritma C4.5 yang memang belum ada yang menganalisa pencapaian hafalan al-Qur'an.

METODE

Tahap pertama penelitian ini adalah studi literatur atau kepastakaan yang dilakukan dengan menelusuri literatur serta menelaahnya untuk menggali teori-teori yang sedang berkembang, mencari metode penelitian yang digunakan terdahulu dan untuk memperoleh orientasi yang ada dalam permasalahan.

Tahap kedua adalah menentukan obyek penelitian yang akan di analisis pencapaian hafalannya. Obyek yang kami pilih adalah program CampTahfizh di Pesantren Tahfizh Daarul Qur'an binaanya KH. Yusuf Mansur di Tangerang. Sengaja obyek ini dipilih karena memiliki santri yang banyak, sehingga banyak data yang dapat dimanfaatkan.

Tahap ketiga adalah menentukan variabel yang akan di proses. Variabel-variabel yang akan diproses datamining ini ditentukan berdasarkan tujuan penelitian. Yaitu catatan santri di awal mengikuti program CampTahfizh, serta hal-hal yang terjadi dalam proses berlangsungnya program tersebut dan hasil yang diperoleh dari program tersebut.

Nilai hafalan sebelum mengikuti program ini akan menjadi tolak ukur dalam perolehan hafalan karena santri telah memiliki pengalaman dalam menghafal al-Qur'an, tambahan waktu diluar halaqohnya akan menjadi tambahan nilai dalam mengukur pencapaiannya. Waktunya pun akan menjadi variabel penting dalam penelitian ini, untuk itu

waktu mulai dan selesainya tepat waktu akan menjadi salah satu variabel ini, karena waktu pun menentukan perolehan hafalannya.

Hal-hal yang termasuk kedalam faktor internal adalah kemampuan intelektual atau kecerdasan (intelegensi), minat, bakat khusus, motivasi untuk berprestasi, sikap, kondisi fisik dan mental, harga diri akademik, dan kemandirian. Kemudian dikemukakan pula hal-hal yang termasuk kedalam faktor eksternal, yaitu lingkungan sekolah, keluarga, dan faktor situasional. (Gunarsa, Y.S. 2000). dari itu kami menerapkan lingkungan bersih, kerajinan mengikuti halaqoh dalam variabel ini, sbegaimana disajikan di tabel 1.

Data mining didefinisikan sebagai proses menemukan dan menggambarkan pola struktural dalam data sebagai alat untuk membantu menjelaskan data dan membuat prediksi dari data tersebut (Varney:2007). Data mining telah menarik banyak perhatian dalam dunia sistem informasi dan dalam masyarakat secara keseluruhan dalam beberapa tahun terakhir, karena ketersediaan luas dalam jumlah besar data dan kebutuhan segera untuk mengubah data tersebut menjadi informasi yang berguna dan pengetahuan.

Informasi dan pengetahuan yang diperoleh dapat digunakan untuk aplikasi mulai dari pasar analisis, deteksi penipuan, dan retensi pelanggan, untuk pengendalian produksi dan ilmu pengetahuan eksplorasi.

Klasifikasi merupakan bagian dari prediksi, dimana nilai yang diprediksi berupa label. Klasifikasi menentukan class atau grup untuk tiap contoh data, input dari model klasifikasi adalah atribut dari contoh data (data samples) dan outputnya adalah class dari data samples itu sendiri, dalam machine learning untuk membangun model klasifikasi digunakan metode supervised learning. (Moghimpour, Ida dan Ebrahimou: 2014).

Algoritma C4.5 diperkenalkan oleh J. Ross Quinlan yang merupakan pengembangan dari algoritma ID3, algoritma tersebut digunakan untuk membentuk pohon keputusan. Pohon keputusan dianggap sebagai salah satu pendekatan yang paling populer, dalam klasifikasi pohon keputusan terdiri dari sebuah node yang membentuk akar, node akar tidak memiliki inputan. Node lain yang bukan sebagai akar tetapi memiliki tepat satu inputan disebut node internal atau test node, sedangkan node lainnya dinamakan daun. Daun mewakili

nilai target yang paling tepat dari salah satu class (Han, J., & Kamber, M:2006).

Pohon keputusan dibuat dengan membagi nilai-nilai atribut menjadi cabang untuk setiap kemungkinan. Cara kerja pohon keputusan yaitu dengan melakukan penelusuran dari akar hingga ke cabang sampai class suatu objek ditemukan, instance diklasifikasikan dengan mengarahkan dari akar pohon sampai ke daun sesuai dengan hasil tes melalui node internal.

Tabel 1. Tabel atribut yang digunakan

NO	Atribut	Nilai
1	Sebelum	<ul style="list-style-type: none"> • 1 - 5 Juz • 5 - 10 Juz • 10 - 15 Juz • 15 - 20 Juz • 25 - 30 Juz • Kurang dari 1 juz
2	Tambah Waktu Diluar	<ul style="list-style-type: none"> • Ya • Mungkin • Tidak
3	Ustad Ada Halaqoh Tidak Saat	<ul style="list-style-type: none"> • 1 - 5 kali • Tidak pernah • 5 - 10 kali • lebih dari 10 kali
4	Halaqoh dimulai dan selesai tepat waktu	<ul style="list-style-type: none"> • Ya • lebih banyak tepat • Tidak • lebih banyak tidak tepat
5	Dimulai setelah bersih	<ul style="list-style-type: none"> • Ya • Mungkin • Tidak
6	Tidak mengiktu halaqoh	
7	Hasil / Setelah	<ul style="list-style-type: none"> • 10 - 15 Juz • 5 - 10 Juz • 1 - 5 Juz • 25 - 30 Juz • 15 - 20 Juz • 20 - 25 Juz

Dari variabel-variabel diatas, kami akan membentuk pohon keputusan sehingga dapat kami terapkan pada aplikasi yang dapat dipergunakan dalam menganalisa hasil di program selanjutnya.

Ada beberapa tahapan dalam membentuk sebuah pohon keputusan dalam algoritma C4.5 (Larose:20015), yaitu:

1. Mempersiapkan data *training*, data *traini* diambil dari riwayat data yang pernah terjadi sebelumnya atau disebut data masa lalu.
2. Menghitung akar dari pohon. Akar akan diambil dari atribut yang akan terpilih, dengan cara menghitung nilai gain dari masing-masing atribut, nilai gain yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai gain dari atribut, hitung dahulu nilai entropy. Untuk menghitung nilai entropy digunakan rumus:

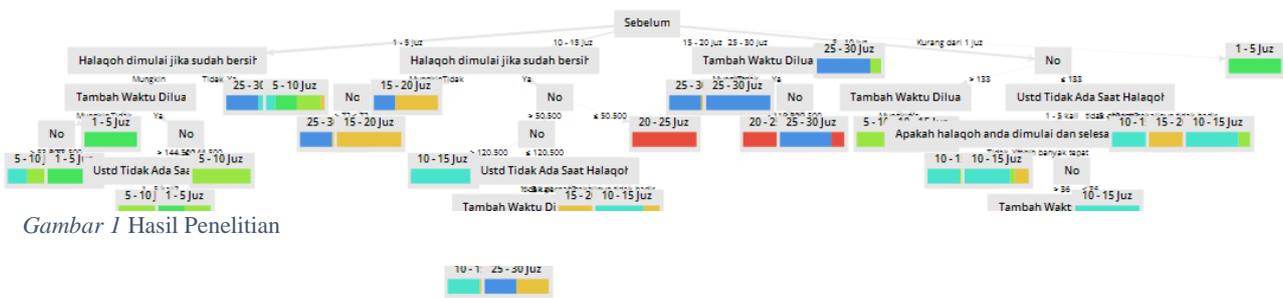
$$Entropy(S) = \sum_{i=1}^n - p_i \log_2 p_i$$

dengan S sebagai himpunan kasus, n jumlah partisi, dan P_i adalah proporsi S_i terhadap S. Setelah entropy didapat, baru kemudian dihitung gain dengan rumus:

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

dengan S sebagai himpunan kasus, A fitur, n jumlah partisi atribut A, $|S_i|$ proporsi S_i terhadap S, $|S|$ jumlah kasus dalam S.

3. Ulangi langkah ke-2 dan langkah ke-3 hingga semua record terpartisi.
4. Proses partisi pohon keputusan akan berhenti saat semua record dalam simpul N mendapat kelas yang sama, tidak ada atribut didalam record yang dipartisi lagi dan tidak ada record dalam cabang yang kosong.



Gambar 1 Hasil Penelitian

HASIL DAN PEMBAHASAN

Tahapan pengolahan awal data melalui datamining. Pengolahan awal data meliputi proses input data ke format yang dibutuhkan, pengelompokan dan penentuan atribut data yaitu sebanyak 659 record. namun tidak semua data dapat digunakan dan tidak semua atribut digunakan karena harus melalui beberapa tahap pengolahan awal data (preparation data). Untuk mendapatkan data yang berkualitas, beberapa teknik yang dilakukan sebagai berikut [9]: Data validation, untuk mengidentifikasi dan menghapus data yang ganjil (outlier/noise), data yang tidak konsisten, dan data yang tidak lengkap (missing value). Data siswa yang ada akan diidentifikasi apakah ada data yang ganjil (outlier/noise), data yang tidak konsisten, dan data yang tidak lengkap (missing value). Data integration and transformation, untuk meningkatkan akurasi dan efisiensi algoritma. Data yang digunakan dalam penulisan ini bernilai kategorikal. Data ditransformasikan kedalam software Rapidminer.

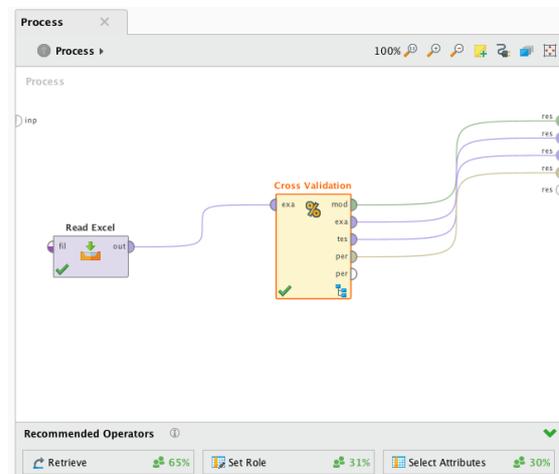
Data size reduction and discretization, untuk memperoleh data set dengan jumlah atribut dan record yang lebih sedikit tetapi bersifat informatif. Pada tahap ini data yang masih berupa nilai numerik dan juga continue ditransformasikan kedalam bentuk kategorikal dan dibuat skala atau interval sehingga menghasilkan range yang lebih kecil sebagai bahan pembelajaran algoritma C4.5.

Tabel 2. Tabel Variabel

NO	Variabel	Atribut
1	Sebelum	Atribut
2	Tambah Waktu Diluar	Atribut
3	Ustad Tidak Ada Saat Halaqoh	Atribut
4	Halaqoh dimulai dan selesai tepat waktu	Atribut
5	Dimulai setelah bersih	Atribut

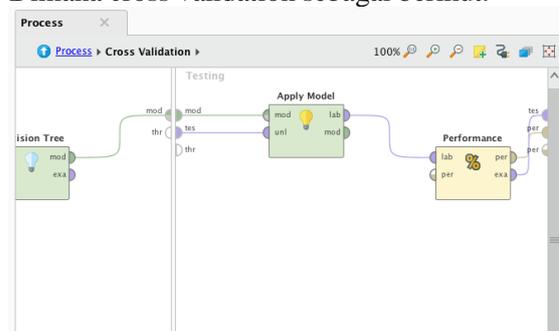
6	Tidak mengiktu halaqoh	Atribut
7	Hasil / Setelah	Label

Dari hasil penelitian, didapatkan pohon keputusan yang digambarkan pada gambar 1, dimana akarnya adalah kolom setelah. Akar menunjukan hasil akhir dari seluruh nilai pada atribut-atribut data yang disajikan pada rapidminer.



Gambar 2 Proses pengolahan menggunakan Rapidminer

Dimana cross validation sebagai berikut:



Gambar 3 Cross Validation

Cross validation kami menggunakan performance confusion matrix, dengan hasil

yang tampilkan accuracy dan kappa pada tabel 3 dan tabel 4.

	true 25 - 30 Juz	true 10 - 15 Juz	true 1 - 5 Juz	true 5 - 10 Juz	true 15 - 20 Juz	true 20 - 25 Juz	class precision
pred. 25 - 30 Juz	13	3	0	2	8	5	41.94%
pred. 10 - 15 Juz	1	32	0	11	8	2	59.26%
pred. 1 - 5 Juz	1	0	10	9	0	0	50.00%
pred. 5 - 10 Juz	0	6	12	10	2	0	33.33%
pred. 15 - 20 Juz	5	2	0	0	3	2	25.00%
pred. 20 - 25 Juz	2	0	0	0	0	0	0.00%
class recall	59.09%	74.42%	45.45%	31.25%	14.29%	0.00%	

Tabel 3. Accuracy Confusion Matrix

Analisa Pencapaian Hafalan Al-Qur'an
Pilih daftar dibawah ini!

Hafalan sebelumnya? 15 - 20 Juz ▼

Tambahan Waktu? Ya ▼

Hasilnya: **Bisa mencapai 25 - 30 Juz**

Gambar 4 Aplikasi penerapan decisin tree

Setelah diolah maka dilakukan teknik pengujian dengan menggunakan confusion matrix dengan jumlah true positive (TP) adalah 13 record diklasifikasikan sebagai 25 mencapai 25-30 Juz, dan False Negative (FN) adalah 18 record diklasifikasikan sebagai belum mencapai 25-30 Juz.

Selanjutnya 9 record untuk True Negative (TN) dikalsifikasikan sebagai tidak mencapai 25-30 Juz, 9 False Positive (FP) diklasifikasikan mencapai 25-30 Juz. Berdasarkan tabel 3.

SIMPULAN DAN SARAN

Dari hasil pengujian menggunakan data yang telah diperoleh. Dilihat dari tabel 3, dari precision dan recall di dapat dengan menggunakan metode ini, santri hanya mendapatkan perolehan hafalan di kategori 10-15 juz.

Oleh sebab itu, penelitian ini harus di lanjutkan ke tahap lebih lanjut lagi, agar mendapatkan hasil yang lebih akurat. Bisa saja ditambahkan beberapa variabel yang mendukung santri dalam menghafal al-Qur'an.

DAFTAR PUSTAKA

- Andriani, Anik. 2012. *PENERAPAN ALGORITMA C4.5 PADA PROGRAM KLASIFIKASI MAHASISWA DROPOUT*. Seminar Nasional Matematika 2012. Hlm 139-147.
- Gorunescu, Florin. 2011. *Data Mining: Concepts and Techniques*. Verlag berlinHeidelberg: Springer.
- Maslihah, Sri. 2011. *STUDI TENTANG HUBUNGAN DUKUNGAN SOSIAL, PENYESUAIAN SOSIAL DI LINGKUNGAN SEKOLAH DAN PRESTASI AKADEMIK SISWA SMPIT ASSYFA BOARDING SCHOOL*. *Jurnal Psikologi Undip Vol. 10, No.2, Oktober 2011*
- Maulana, Arif. 2006. *Pengembangan Kurikulum Pelatihan Magang di STM Nasional Semarang: Suatu Studi Berdasarkan Dunia Usaha*. Tesis tidak diterbitkan. Semarang: PPS UNNES
- Nugroho, Yusuf Sulisty. 2014. *PENERAPAN ALGORITMA C4.5 UNTUK KLASIFIKASI PREDIKAT KELULUSAN MAHASISWA FAKULTAS KOMUNIKASI DAN INFORMATIKA UNIVERSITAS MUHAMMADIYAH SURAKARTA*. Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) 2014. Yogyakarta, 15 November 2014.
- Utami, Lila Dini & Wahono, Romi Satria. 2015. *Integrasi Metode Information Gain Untuk Seleksi Fitur dan Adaboost Untuk Mengurangi Bias Pada Analisis Sentimen Review Restoran Menggunakan Algoritma Naïve Bayes*. *Journal of Intelligent Systems*, Vol. 1, No. 2, December 2015. hlm 120-126.
- Puspita, Ari & Wahyudi, Mochamad. 2015. *ALGORITMA C4.5 BERBASIS DECISION TREE UNTUK PREDIKSI KELAHIRAN BAYI PREMATUR*. Konferensi Nasional Ilmu Pengetahuan dan Teknologi (KNIT) 2015 Bekasi, hlm 107-112