

PREDIKSI PENCAPAIAN HAFALAN AL-QUR'AN MENGGUNAKAN METODE C4.5 BERBASIS PSO

Muchlis, Said Mirza Pahlevi

Program Pascasarjana, STMIK Nusamandiri, Jakarta Pusat, Jl. Kramat Raya Nomor 18
muchlis.re@gmail.com.

Abstrak

Beberapa tahun yang lalu sampai sekarang, banyak universitas negeri maupun swasta memberikan beasiswa untuk siswa yang memiliki hafalan al-Qur'an yang baik. Untuk itu, memprediksi hafalan al-Qur'an lebih awal akan memberikan kesempatan lebih bagi lembaga untuk mempersiapkan siswa-siswa agar dapat mencapai target. Penelitian ini menggunakan data nilai akademik, jumlah hafalan dan kuisioner pada Pesantren Tahfiz Daarul Qur'an, yaitu data siswa tahun pelajaran 2014/2015. Penelitian menggunakan algoritma klasifikasi C4.5, C4.5 berbasis PSO dan k-NN. Untuk mengevaluasi hasil penelitian ketiga algoritma tersebut, peneliti menggunakan metode pengujian Confusion Matrix. Nilai akurasi yang didapat untuk algoritma tertinggi adalah algoritma C4.5 berbasis PSO sebesar 89.09%, untuk algoritma C4.5 memiliki akurasi sebesar 82.64% dan algoritma k-NN memiliki akurasi 66.00%. Hasil uji beda menggunakan Anova telah menghasilkan nilai dibawah nilai alpa, jadi dari ketiga algoritma signifikan perbedaannya. Dapat diambil kesimpulan bahwa algoritma C4.5 berbasis PSO adalah yang terbaik.

Kata kunci: *C4.5, PSO, hafalan, Qur'an, prediksi*

Abstract

Several years ago until now, many universities provide scholarships for students who have a good recitation of the Qur'an. Therefore, predicting the recitation of the Qur'an earlier will provide more opportunities for the institution to prepare the students to reach the target. This study uses data of academic value, the number of memorization and questionnaire at Pesantren Tahfiz Daarul Qur'an, it is a data the student in 2014/2015 academic year. The study used classification algorithm C4.5, C4.5 with PSO and k-NN. To evaluate the research results of the three algorithms, the researchers used Confusion Matrix. The accuracy value obtained for the highest algorithm is the C4.5 with PSO of 89.09%, for the algorithm C4.5 has an accuracy of 82.64% and the k-NN algorithm has an accuracy of 66.00%. Differential test shown that result the value is under the alpha value, Therefore the differens of three algoritms is significant. It can be concluded that the C4.5 with PSO algorithm is the best.

Keywords : *C4.5, PSO, memorization, Qur'an, predictions*

PENDAHULUAN

Pesantren memiliki pengaruh yang sangat besar dalam perkembangan pendidikan di Indonesia. Pesantren awalnya hanya lembaga pendidikan yang mengajarkan pendidikan agama. Tetapi pada perkembangannya, pesantren memperluas materi pembelajarannya ke pembelajaran umum (Dofier,1983).

Santri merupakan objek dari sebuah pesantren dan menjadi aspek penting dalam

evaluasi keberhasilan penyelenggaraan pembelajaran pada suatu pesantren. Evaluasi santri terhadap kualitas, prestasi, akhlak, kebiasaan dan kompetensi lulusan.

Pesantren Tahfiz Daarul Qur'an adalah lembaga pendidikan yang fokus kepada pendidikan menghafal al-Qur'an, atau disebut tahfiz. Penentuan kelulusan santri didasarkan kepada seberapa baik hafalan yang dimiliki oleh santri selama menempuh pembelajaran di

pesantren, yaitu memiliki target 30 juz hingga santri dapat dinyatakan lulus.

Santri yang memiliki hafalan 30 juz memiliki kesempatan khusus untuk masuk ke perguruan tinggi negeri atau swasta, dalam negeri maupun luar negeri.

Oleh karena itu, santri perlu dipersiapkan agar dapat memperoleh hafalan al-Qur'an sampai juz 30. Penentuan klasifikasi santri-santri harus dilakukan, agar pesantren dapat menentukan langkah-langkah yang harus dilakukan terhadap santri-santri yang masuk kedalam kategori klasifikasi kurang.

Penelitian-penelitian dalam menentukan klasifikasi banyak menggunakan algoritma C4.5 dan k-NN. Seperti (Kamagi & Hansun, 2014) yang melakukan penelitian tentang prediksi kelulusan serta pencapaian nilai akhir semester. (MALIK & FAID, 2016) yang memprediksi tingkat kelulusan siswa, (Aidi, 2016; Witno, 2017; Panoto, Utami, 2017) yang memprediksi kelulusan siswa menggunakan C4.5 dan k-NN mendapatkan akurasi yang tinggi.

(Hijriana, 2008; Ridwan, 2017) yang memprediksi siswa menggunakan C4.5 dengan akurasi 72.47%, bahkan sampai 99.10% di penelitian Hijriana. Penelitian tentang pencapaian hafalan al-Qur'an pernah dilakukan (Muchlis & Taufiqurochman, 2017), tetapi mendapatkan akurasi yang sangat kecil. Oleh karena itu, perlu dilakukan penelitian lanjutan untuk memprediksi hafalan al-Qur'an dengan mengklasifikasi kategori hafalan santri..

LANDASAN TEORI

Klasifikasi yaitu proses pencarian pola yang membedakan kelas atau konsep data. Pola ini digunakan untuk memprediksi label kelas objek yang labelnya tidak diketahui (Jiawei, Kamber, Han, Kamber, & Pei, 2012).

Decision Tree struktur pohon dimana masing-masing simpul internal menunjukkan sebuah tes pada atribut, masing-masing cabang mewakili hasil pengujian, dan setiap simpul daun memegang label kelas (Jiawei et al., 2012).

k-Nearest Neighbor (k-NN) melakukan prediksi dengan membandingkan tupel test yang diberikan dengan tupel training yang serupa dengannya (Jiawei et al., 2012).

Particle Swarm Optimization (PSO) adalah satu teknik komputasi evolusioner yang dikembangkan yang terinspirasi dari kawanan

burung yang mencari makanan. PSO adalah algoritma iterasi yang berbasis populasi (Jun, Cho, & Lee, 2017).

Penelitian yang memprediksi pencapaian hasil pembelajaran akademik menggunakan algoritma C4.5 atau k-NN cukup banyak dan mendapatkan nilai akurasi yang tinggi.

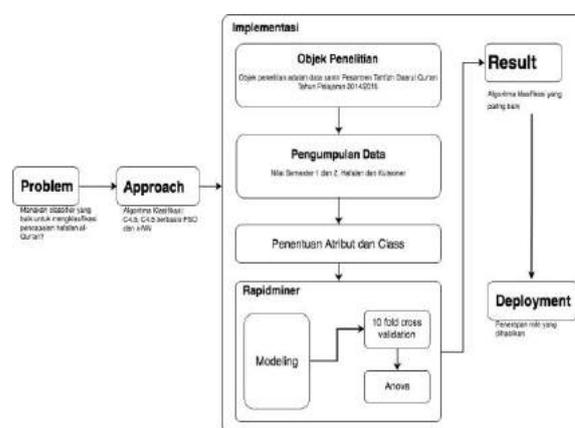
Objek penelitian adalah Pesantren Tahfizh Daarul Qur'an yang merupakan salah satu lembaga pendidikan yang memiliki karakteristik dalam mengkhususkan pembelajarannya pada bisang tahfizhul Qur'an.

Pesantren adalah pendidikan dan pengajaran agama islam dari seorang atau beberapa guru yang disebut kyai dan ustadz dan santri sebagai murid (Tim Direktorat Jenderal Kelembagaan Agama Islam, 2003).

Hafalan atau menghafal adalah berusaha meresapkan kedalam pikiran agar selalu ingat (Tim Prima Pena). Dalam Bahasa arab, bias disebut al-Hafizh yang artinya menjaga, memelihara atau menghafalkan (Munawir, 1997).

METODOLOGI PENELITIAN

Penelitian menggunakan metode CRISP-DM (Cross Industry Standard Process for Data Mining) yang dimasukkan kedalam kerangka kerja penelitian yang ditunjukkan dalam Gambar 1.



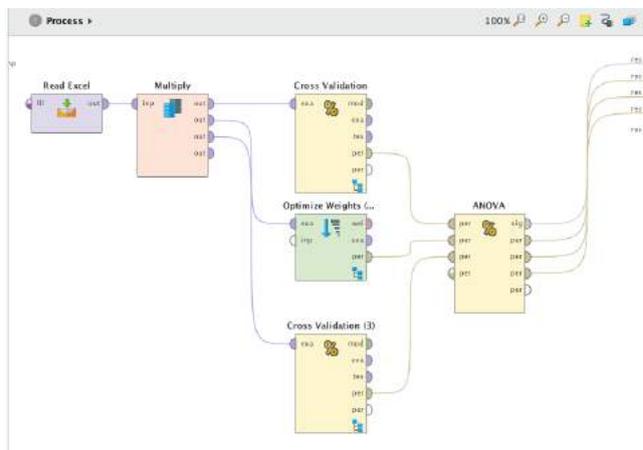
Gambar 1. Kerangka kerja penelitian

- Problem. Yaitu mencari classifier yang baik untuk mengklasifikasi pencapaian hafalan al-Qur'an.
- Approach. Algoritma klasifikasi C4.5, C4.5 berbasis PSO dan k-NN akan digunakan dalam penelitian ini.

c. Implementasi. Data yang digunakan adalah data santri Pesantren Tahfizh Daarul Qur'an tahun pelajaran 2014/2015 dengan jumlah 109 santri. Di langkah ini akan ditentukan objek penelitian, pengumpulan data, penentuan atribut dan diproses menggunakan aplikasi Rapidminer.

HASIL DAN PEMBAHASAN

Model digambarkan pada Gambar 2 dengan 109 jumlah data santri kelas 11 tahun pelajaran 2014/2015 yang terdiri dari nilai akademik, nilai tahfizh dan kuisisioner.



Gambar 2. Desain proses perbandingan algoritma C4.5, C4.5 berbasis PSO dan k-NN.

Hafalan santri adalah banyaknya juz yang dihafal, dari juz 1 sampai juz 30. Oleh karena itu dikonversikan kedalam 3 kategori Tabel 1.

Tabel 1. Tabel kategori perolehan hafalan

| Jumlah Juz | Kategori |
|-------------------|-------------|
| 1 Juz s.d 10 Juz | Kurang |
| 11 Juz s.d 20 Juz | Baik |
| 21 Juz s.d 30 Juz | Sangat Baik |

Dari model Gambar 2 yang diproses menggunakan aplikasi Rapidminer dengan pengujian menggunakan *Cross Validation* yang melakukan proses *10-fold cross validation* ini untuk ketiga algoritma yang digunakan.

Tabel 2. Confusion Matrix Algoritma C4.5

accuracy: 82.64% +/- 9.42% (mikro: 82.57%)

| | true SANGAT BAIK | true BAIK | true KURANG | class precision |
|-------------------|------------------|-----------|-------------|-----------------|
| pred. SANGAT BAIK | 29 | 8 | 0 | 78.38% |
| pred. BAIK | 3 | 19 | 4 | 73.08% |
| pred. KURANG | 0 | 4 | 42 | 91.30% |
| class recall | 90.62% | 61.29% | 91.30% | |

Tabel 3. Confusion Matrix Algoritma C4.5 berbasis PSO

accuracy: 89.09% +/- 6.80% (mikro: 88.99%)

| | true SANGAT BAIK | true BAIK | true KURANG | class precision |
|-------------------|------------------|-----------|-------------|-----------------|
| pred. SANGAT BAIK | 30 | 5 | 0 | 85.71% |
| pred. BAIK | 2 | 22 | 1 | 88.00% |
| pred. KURANG | 0 | 4 | 45 | 91.84% |
| class recall | 93.75% | 70.97% | 97.83% | |

Tabel 4. Confusion Matrix Algoritma k-NN

accuracy: 66.00% +/- 10.94% (mikro: 66.06%)

| | true SANGAT BAIK | true BAIK | true KURANG | class precision |
|-------------------|------------------|-----------|-------------|-----------------|
| pred. SANGAT BAIK | 24 | 12 | 3 | 61.54% |
| pred. BAIK | 7 | 14 | 9 | 46.67% |
| pred. KURANG | 1 | 5 | 34 | 85.00% |
| class recall | 75.00% | 45.16% | 73.91% | |

Berdasarkan tabel confusion matrix dari Tabel 2, 3 dan 4 didapatkan akurasi 82.64% untuk algoritma C4.5, akurasi 89.09% untuk algoritma C4.5 berbasis PSO dan 66% untuk akurasi k-NN.

Berdasarkan akurasi yang didapat, dilakukan uji test menggunakan Anova diperoleh bahwa rata-rata akurasi dari ketiga algoritma adalah signifikan perbedaannya dengan nilai 0.000 yang ditunjukkan pada Tabel 5.

Tabel 5. Anova untuk C4.5, C4.5 berbasis PSO dan k-NN

| Source | Square Sums | DF | Mean Squares | F | Prob |
|-----------|-------------|----|--------------|--------|-------|
| Between | 0.284 | 2 | 0.142 | 16.721 | 0.000 |
| Residuals | 0.229 | 27 | 0.008 | | |
| Total | 0.513 | 29 | | | |

KESIMPULAN

Dari penelitian ini diperoleh bahwa algoritma C4.5 yang berbasis PSO memperoleh akurasi yang paling tinggi dengan nilai akurasi mencapai 89.09%. Klasifikasi pencapaian al-Qur'an menggunakan algoritma C4.5 berbasis PSO dengan data ini lebih baik dalam mengklasifikasi kategori Kurang berdasarkan hasil pengukuran F1 Measure dengan nilai 92.63%.

DAFTAR PUSTAKA

Hijriana, N. (2008). Penerapan metode decision tree algoritma c4.5 untuk klasifikasi mahasiswa berprestasi, 39–43.

- Jiawei, H., Kamber, M., Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. San Francisco, CA, itd: Morgan Kaufmann.
<http://doi.org/10.1016/B978-0-12-381479-1.00001-0>
- Jun, C., Cho, Y., & Lee, H. (2017). Improving Tree-based Classification Rules Using a Particle Swarm Optimization.
- Kamagi, D. H., & Hansun, S. (2014). Implementasi Data Mining dengan Algoritma C4 . 5 untuk Memprediksi Tingkat Kelulusan Mahasiswa. *ULTIMATICS*, Vol. VI, No. 1 | Juni 2014, VI(1), 15–20.
- MALIK, K., & FAID, M. (2016). Prediksi Prestasi Siswa Smp Nurul Jadid Menggunakan Algoritma C4.5. *Nusantara Journal of Computers and Its Applications*, 2(4), 2–5.
- Muchlis, & Taufiqurochman. (2017). Analisa pencapaian hafalan al- qur'an dalam camptahfizh dengan algoritma c.45 dan penerapannya. *SEMNASTEK UMJ 2017*, (November), 1–2.
- Munawir, A.W. (1997) *Almunawir Kamus Bahasa Arab-Indonesia*. Surabaya: Pustaka Progresif.
- Ridwan, M. (2017). Sistem Rekomendasi Proses Kelulusan Mahasiswa Berbasis Algoritma Klasifikasi C4.5. *Jurnal Ilmiah Informatika*, 2(1), 105–111.
- Tim Direktorat Jendral Kelembagaan Agama Islam. (2003). *Pola Pembelajaran di Pesantren*. Jakarta: Departemen Agama.