

Random Forest Classifier Dengan Grid Search Untuk Klasifikasi Hasil Psikotes Pada Rekrutmen Dosen dan Tenaga Kependidikan

Nico Pratama^{1,*}, Endang Setyati², Joan Santoso³, Amanda Pasca Rini⁴

^{1,2,3}Magister Teknologi Informasi, Pascasarjana Teknologi Informasi,
Insitut Sains dan Teknologi Terpadu Surabaya

⁴Magister Psikologi, Fakultas Psikologi, Universitas 17 Agustus 1945 Surabaya

ABSTRACT

Effective selection of educational staff and lecturers requires appropriate evaluation techniques to assess the abilities and potential of candidates. This research recommends using the Random Forest Classifier to classify psychological test results, with optimization using Grid Search and validation via K-Fold Cross Validation. This approach was chosen because of its ability to handle complex and heterogeneous data. Grid Search is used to search for the best parameter combination that improves model performance, while K-Fold Cross Validation ensures generalization reliability. This research used 616 data, consisting of 307 lecturers' and 309 psychological test results educational staff psychological test results. Data is divided into training (80%) and testing (20%) data. Classification report analysis shows an accuracy of 85% for the education staff dataset and 82% for the lecturer dataset.

Keywords: *Random Forest Classifier; Grid Research; Psychological Test Results; K-Fold Cross Validation*

ABSTRAK

Seleksi tenaga kependidikan dan dosen yang efektif memerlukan teknik evaluasi yang tepat untuk menilai kemampuan serta potensi kandidat. Penelitian ini merekomendasikan penggunaan Random Forest Classifier untuk mengklasifikasikan hasil psikotes, dengan optimasi menggunakan Grid Search dan validasi melalui K-Fold Cross Validation. Pendekatan ini dipilih karena kemampuannya dalam menangani data yang kompleks dan heterogen. Grid Search digunakan untuk mencari kombinasi parameter terbaik yang meningkatkan kinerja model, sementara K-Fold Cross Validation memastikan keandalan generalisasi. Penelitian ini menggunakan 616 data, terdiri dari 307 hasil psikotes dosen dan 309 hasil psikotes tenaga kependidikan. Data dibagi menjadi data pelatihan (80%) dan pengujian (20%). Analisis classification report menunjukkan akurasi sebesar 85% untuk dataset tenaga kependidikan dan 82% untuk dataset dosen.

Kata Kunci: *Random Forest Classifier; Grid Research; Hasil Psikotes; K-Fold Cross Validation*

1. Pendahuluan

Kebutuhan akan rekrutmen yang Efisien, Universitas atau institusi pendidikan seringkali menghadapi tantangan dalam merekrut dosen dan staf administratif (tenaga kependidikan) baru. Dalam proses rekrutmen, psikotes sering digunakan untuk mengukur kualitas psikologis calon karyawan. Psikotes dapat menjadi proses yang kompleks karena melibatkan berbagai jenis tes dan evaluasi yang harus dianalisis secara holistik. Tingkat Kompetisi yang Tinggi Banyak institusi pendidikan bersaing untuk menarik dan merekrut dosen dan staf administratif terbaik [1].

Menentukan kualifikasi seseorang sebelumnya memerlukan waktu yang cukup lama bagi seorang psikolog atau pakar. Mereka harus menganalisis jawaban dengan cermat dan mengolah keluaran tes menggunakan metode manual. Namun, dengan kemajuan teknologi saat ini, proses ini menjadi jauh lebih mudah [2].

Penggunaan *Random Forest Classifier* dengan *Grid Search* dalam klasifikasi hasil psikotes pada rekrutmen dosen dan tenaga kependidikan memiliki keunggulan yang sesuai dengan karakteristik algoritma tersebut. *Random Forest* mampu mengatasi kompleksitas data dan *overfitting* melalui pembangunan pohon keputusan secara acak, sementara *Grid Search* memungkinkan optimasi parameter yang sistematis untuk meningkatkan kinerja model [3]. Dengan menggunakan metode *Random Forest* yang dilengkapi dengan *Grid Search*, institusi pendidikan dapat menunjukkan kemampuan adaptasi terhadap perkembangan teknologi yang dapat memberikan keunggulan kompetitif. Proses rekrutmen yang terorganisir, efisien, dan adil dapat meningkatkan reputasi institusi di mata calon karyawan, dan sekaligus membantu mempertahankan dan menarik individu berbakat. Tujuan dari pembuatan sistem ini adalah untuk meningkatkan efisiensi dan objektivitas dalam proses rekrutmen dosen dan tenaga kependidikan melalui penggunaan teknologi *Machine Learning*. Dengan menerapkan algoritma *Random Forest Classifier* yang telah dioptimalkan dengan *Grid Search*, sistem ini bertujuan untuk meningkatkan akurasi, mengurangi bias subjektif, dan meningkatkan efisiensi proses seleksi [4].

2. Tinjauan Pustaka

Tinjauan pustaka dalam penelitian ini mencakup kajian mengenai penggunaan *machine learning* dalam klasifikasi hasil psikotes, optimasi model menggunakan *Grid Search*, serta evaluasi model melalui *k-fold cross validation*. *Machine learning* telah banyak digunakan dalam berbagai bidang untuk meningkatkan akurasi prediksi dan pengambilan keputusan berbasis data. Algoritma *Random Forest Classifier*, yang dikembangkan oleh [5], dikenal karena kemampuannya dalam menangani data yang kompleks dan heterogen

serta menghasilkan model yang *robust* dan interpretatif. Beberapa penelitian sebelumnya telah menunjukkan keefektifan *Random Forest* dalam berbagai aplikasi, termasuk klasifikasi dan regresi [6]. Untuk meningkatkan kinerja model, teknik *Grid Search* digunakan untuk mengoptimalkan hiperparameter model, seperti yang dijelaskan oleh [7], yang menunjukkan bahwa optimasi parameter yang tepat dapat secara signifikan meningkatkan performa model *machine learning*. Evaluasi model menggunakan *k-fold cross validation*, seperti yang diuraikan oleh [8], memberikan estimasi performa yang lebih akurat dan mengurangi risiko *overfitting* dengan membagi data pelatihan menjadi beberapa subset. Dalam konteks klasifikasi hasil psikotes, penelitian sebelumnya oleh [9] menunjukkan bahwa penggunaan metode *machine learning* dapat membantu dalam meningkatkan keandalan evaluasi psikotes dan mendukung pengambilan keputusan dalam proses rekrutmen. Penelitian ini berusaha untuk mengintegrasikan metode-metode tersebut untuk menghasilkan model klasifikasi yang optimal dan handal dalam konteks rekrutmen dosen dan tenaga kependidikan.

Tinjauan pustaka ini juga mencakup studi tentang penerapan algoritma *machine learning* dalam rekrutmen dan manajemen sumber daya manusia. Misalnya, penelitian oleh [10] yang mengkaji penggunaan model prediktif untuk mendukung proses rekrutmen dan seleksi karyawan, menunjukkan bahwa algoritma *machine learning* dapat meningkatkan efisiensi dan efektivitas dalam memilih kandidat yang tepat. Selanjutnya, penelitian oleh [1] membahas tentang klasifikasi hasil psikotes menggunakan algoritma *Random Forest* dan menemukan bahwa algoritma ini dapat mengolah data psikotes dengan baik dan memberikan hasil yang akurat. *Grid Search*, seperti yang dijelaskan oleh [7], adalah metode pencarian sistematis untuk mengoptimalkan parameter yang paling mempengaruhi performa model, memastikan bahwa model beroperasi pada tingkat kinerja terbaiknya. Evaluasi menggunakan *k-fold cross validation* juga telah diterapkan secara luas dalam penelitian *machine learning* untuk memberikan estimasi yang lebih stabil dan mengurangi bias yang mungkin muncul dari pembagian data yang tidak seimbang [11]. Dalam penelitian ini, kombinasi dari *Random Forest Classifier*, optimasi *Grid Search*, dan evaluasi *k-fold cross validation* diterapkan untuk menghasilkan model klasifikasi yang akurat dan dapat diandalkan dalam konteks rekrutmen dosen dan tenaga kependidikan, memberikan kontribusi baru pada literatur tentang penggunaan *machine learning* dalam pengolahan data psikotes [6].

3. Metode

Metode penelitian ini menggunakan pendekatan kuantitatif dengan desain eksperimen untuk mengembangkan dan mengevaluasi model klasifikasi hasil psikotes dalam rekrutmen dosen dan tenaga kependidikan [12], [13]. Data yang digunakan dalam penelitian ini terdiri dari 616 data hasil psikotes, dengan 307 data calon dosen dan 309 data calon tenaga kependidikan yang diperoleh dari institusi pendidikan. Data mentah ini kemudian melalui tahap *preprocessing* yang mencakup pembersihan data, normalisasi, dan ekstraksi fitur penting. Data dibagi menjadi dua bagian, yaitu 80% untuk pelatihan dan 20% untuk pengujian. Selanjutnya, data pelatihan dibagi lebih lanjut menggunakan metode *k-fold cross validation*, di mana data dibagi menjadi 10 *fold* untuk memastikan validasi yang lebih baik dan generalisasi model.

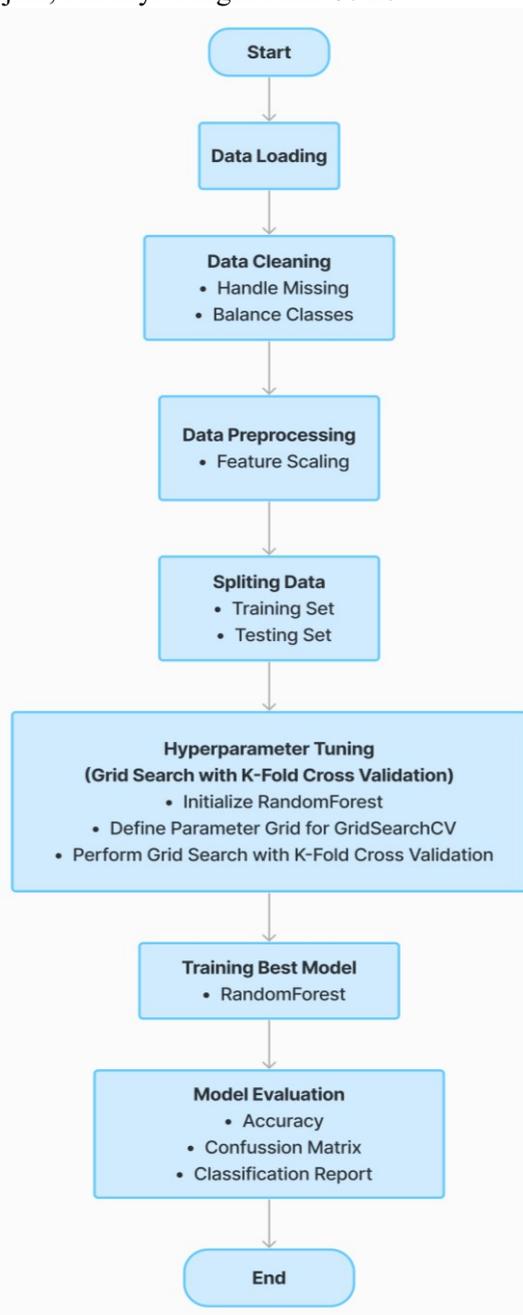
Algoritma *Random Forest Classifier* digunakan untuk membangun model, dengan optimasi hiperparameter melalui teknik *Grid Search* yang mencakup parameter *n_estimators*, *max_depth*, dan *max_features* [14], [15]. Model dilatih dengan data pelatihan yang sudah dioptimalkan dan kinerjanya dievaluasi menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*. Evaluasi dilakukan dengan menghasilkan *classification report* dan *confusion matrix* untuk data pengujian. Hasil evaluasi dibandingkan dengan penelitian sebelumnya untuk mengidentifikasi peningkatan performa dan keandalan model dalam klasifikasi hasil psikotes. Penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam proses rekrutmen dosen dan tenaga kependidikan dengan menyediakan model klasifikasi yang akurat dan dapat diandalkan.

3.1. Desain Penelitian

Studi ini mengadopsi pendekatan eksperimental dengan menggunakan teknik *Machine Learning* untuk mengklasifikasikan hasil psikotes dalam proses rekrutmen dosen dan tenaga kependidikan. Metode *K-fold cross-validation* digunakan untuk memvalidasi model [10].

Proses penelitian pada Gambar 1 dimulai dengan langkah-langkah yang terstruktur untuk memastikan bahwa model yang dibangun memiliki akurasi dan keandalan yang tinggi. Pertama, dilakukan proses data loading, yaitu memuat data dari file Excel yang berisi hasil psikotes calon dosen dan tenaga kependidikan. Setelah data dimuat, langkah berikutnya adalah data *cleaning*, di mana nilai yang hilang dihapus dan kelas data diseimbangkan untuk mengatasi masalah ketidakseimbangan kelas. Selanjutnya, dilakukan data *preprocessing*, yaitu *scaling* pada fitur-fitur data untuk memastikan bahwa semua fitur berada dalam skala yang sama. Setelah data siap, proses *splitting* data dilakukan

dengan membagi data menjadi set pelatihan dan pengujian, biasanya dengan rasio 80:20.



Gambar 1. Diagram Alur Penelitian

Pada tahap *hyperparameter tuning*, dilakukan optimasi menggunakan *Grid Search* yang dikombinasikan dengan *k-fold cross validation*. Pertama, *Random Forest Classifier* diinisialisasi tanpa melatih model. Kemudian, parameter *grid* untuk *GridSearchCV* didefinisikan, mencakup kombinasi parameter yang akan diuji. *Grid Search* dengan *k-fold cross validation* dilakukan untuk menemukan kombinasi parameter terbaik dengan cara melatih dan memvalidasi model menggunakan *k-fold cross validation*. Setelah parameter terbaik ditemukan, *training best model* dilakukan dengan

melatih model *Random Forest Classifier* terbaik pada data pelatihan.

Setelah model dilatih, langkah berikutnya adalah model *evaluation*, di mana performa model dievaluasi menggunakan metrik seperti akurasi, *confusion matrix*, dan *classification report*. *Feature importance visualization* dilakukan untuk menampilkan pentingnya masing-masing fitur dalam model, memberikan wawasan tentang fitur mana yang paling berpengaruh. Selain itu, *tree visualization* dilakukan dengan memvisualisasikan salah satu pohon keputusan dalam hutan, untuk memberikan gambaran tentang bagaimana model membuat keputusan. Proses penelitian ini diakhiri dengan *end*, menandakan bahwa semua tahap telah selesai dilakukan.

3.2. Data Penelitian

Data penelitian terdiri dari hasil psikotes dari 616 responden yang telah melamar untuk posisi dosen dan tenaga kependidikan di Universitas 17 Agustus 1945 Surabaya. Data ini terdiri dari berbagai fitur psikometrik dan label klasifikasi yang menunjukkan apakah kandidat lulus atau gagal dalam seleksi

Dataset yang digunakan dalam penelitian ini diperoleh dari hasil psikotes yang dilakukan pada calon dosen dan tenaga kependidikan. Data diperoleh dari Unit Pelayanan dan Konsultasi Psikologi (UPKP) Universitas 17 Agustus 1945 Surabaya. Dataset yang digunakan dalam penelitian merupakan data teks hasil psikotes sebanyak 616 data. Dataset disimpan dalam bentuk *file Microsoft Excel*. Disarankan, dipertimbangkan, dan tidak disarankan adalah tiga kategori label kelas yang digunakan untuk klasifikasi.

Ada elemen dari masing-masing label kelas klasifikasi yang mempengaruhi hasil psikotes. Hasil psikotes dosen menunjukkan empat aspek: Inteligensi, Sikap Kerja, Kepribadian, dan Kepemimpinan. Hasil psikotes tenaga kependidikan menunjukkan tiga aspek: Inteligensi, Sikap Kerja, dan Kepribadian.

3.3. Preprocessing

Data psikotes diproses untuk menghilangkan nilai yang hilang dan disesuaikan skala nilainya jika diperlukan. Selanjutnya, data dipisahkan menjadi subset pelatihan dan pengujian untuk melatih dan menguji model. Metode *K-fold cross-validation* digunakan untuk membagi data pelatihan menjadi k subset, di mana setiap subset digunakan secara bergantian sebagai data validasi saat model dilatih pada $k-1$ subset lainnya.

Selanjutnya, data diubah untuk menjadi lebih representatif. Dalam penelitian ini, *preprocessing* terdiri dari:

- Data *reduction*, yaitu menghilangkan data yang memiliki nama yang sama atau menduplikasi.

- Pemformatan data dari format *excel* ke *csv*.
- Data *cleansing*, *balancing*, dan standarisasi.
- Transformasi data, yang berarti mengubah data menjadi bentuk yang lebih sesuai.

3.4. Model Machine Learning dan Grid Search

Algoritma *Random Forest Classifier* digunakan untuk mengklasifikasikan hasil psikotes. Teknik *Grid Search* digunakan untuk mencari kombinasi parameter optimal untuk meningkatkan kinerja model.

Algoritma ini berupa kombinasi dari beberapa *tree predictors* atau bisa disebut *decision trees* dimana setiap *tree* bergantung pada nilai *random vector* yang dijadikan sampel secara bebas dan merata pada semua *tree* dalam *forest* tersebut. Hasil prediksi dari *Random Forest* didapatkan melalui hasil terbanyak dari setiap individual *decision tree* (voting untuk klasifikasi dan rata-rata untuk *regresi*). Untuk *Random Forest* yang terdiri dari N trees dirumuskan sebagai:

$$l(y) = \operatorname{argmax}_c \left(\sum_{n=1}^N I_{h_n(y)=c} \right) \quad (1)$$

Dimana I adalah fungsi indikator dan h adalah *tree* ke- n dari *Random Forest*.

Algoritma Klasifikasi *Random Forest* adalah sebagai berikut:

- Langkah 1: *Load dan Preprocess Data*
 - Baca data dari file Excel dan tangani missing values.
 - Lakukan *balancing* kelas (*opsional*) untuk memastikan jumlah data seimbang.
 - Pisahkan fitur (X) dan label (y).
- Langkah 2: *Split Data dan Standardize Features*
 - Bagi data menjadi set pelatihan dan pengujian dengan rasio 80-20.
 - Standarisasi fitur menggunakan `'StandardScaler'`.
- Langkah 3: *Define Model dan Parameter Grid*
 - Inialisasi `'Random Forest Classifier'`.
 - Definisikan parameter *grid* untuk pencarian *grid*.
- Langkah 4: *Perform Grid Search dengan Cross-Validation*
 - Inialisasi `'Grid Search CV'` dengan *10-fold cross-validation*.
 - Latih model menggunakan `'Grid Search CV'`.
- Langkah 5: *Evaluate Model*
 - Dapatkan estimator terbaik dari pencarian *grid*.
 - Evaluasi model terbaik menggunakan data pengujian dan cetak laporan klasifikasi.

Dalam penelitian ini, klasifikasi *random forest* digunakan dengan *grid search*. Setelah melakukan *Grid Search* dengan *K-Fold Cross Validation* pada *Random*

Forest Classifier, kita mendapatkan parameter terbaik yang menghasilkan model dengan kinerja terbaik pada data pelatihan. Berikut adalah penjelasan dan parameter yang digunakan pada penelitian.

- *n_estimators*: 200
Menentukan bahwa 200 pohon digunakan dalam model. Semakin banyak pohon, biasanya model menjadi lebih stabil dan akurat.
- *criterion*: *entropy*
Menggunakan *entropy* sebagai kriteria pemisahan, yang mengukur *impurity* dalam *node*. *Entropy* cenderung lebih baik dalam memisahkan data dibandingkan *Gini* dalam beberapa kasus.
- *min_samples_split*: 2
Node akan dibagi jika memiliki minimal 2 sampel. Nilai yang lebih rendah dapat menyebabkan pohon yang lebih kompleks.
- *min_samples_leaf*: 1
Node akhir (*leaf*) harus memiliki minimal 1 sampel. Nilai yang lebih rendah dapat menyebabkan pohon yang lebih kompleks.
- *max_features*: *'sqrt'*
Untuk setiap pembagian, hanya dipertimbangkan akar kuadrat dari jumlah total fitur. Hal ini mengurangi korelasi antar fitur.
- *max_depth*: 10
Kedalaman maksimal pohon adalah 10. Ini mencegah *overfitting* dengan mengontrol kompleksitas model.
- *random_state*: 42
Seed untuk pengacakan, memastikan hasil yang dapat direplikasi.

Menggunakan parameter terbaik ini, model *Random Forest* akan lebih akurat dan efisien pada data yang diberikan

3.4. Split Test

Pada penelitian ini, tes dibagi menjadi dua bagian: *training* data dan tes data. Rasio *training* dan tes data dalam penelitian ini adalah 80:20. Selain itu, *cross-validation* digunakan dalam *training* data untuk meningkatkan validitas model. *K-Fold Cross* validasi, yang memiliki 10 kali validasi, digunakan sebagai *cross validation* [2].

3.5. Evaluasi

Model *Random Forest Classifier* dengan *Grid Search* berhasil mencapai akurasi sebesar 85% untuk dataset tenaga kependidikan dan 82% untuk dataset dosen, dalam mengklasifikasikan hasil psikotes. Evaluasi oleh para ahli menunjukkan konsistensi yang baik dengan kriteria psikologis, dengan model berhasil mengidentifikasi kandidat yang sesuai dengan standar

psikometrik yang berlaku. Meskipun demikian, para ahli menekankan pentingnya mempertimbangkan juga faktor-faktor non-psikometrik dalam proses rekrutmen.

4. Hasil dan Pembahasan

Data yang telah dikumpulkan melalui hasil psikotes dosen dan tenaga kependidikan di Universitas 17 Agustus 1945 Surabaya terbagi menjadi 3 kategori yaitu Disarankan, Dipertimbangkan dan Tidak Dipertimbangkan. Data keseluruhan yang dikumpulkan ada sebanyak 616 data set yang terdiri dari data hasil psikotes Dosen sebanyak 307 data dan hasil psikotes tenaga kependidikan sebanyak 309 data.

Hasil dari data tersebut kemudian akan dilakukan *training* data. Kriteria data yang dihasilkan dari data dosen memiliki total 24 kriteria penilaian yang mana dikategorikan menjadi 4 aspek. Sementara pada data tendik, dihasilkan total 21 kriteria penilaian yang dikategorikan menjadi 3 aspek.

Dari hasil *training* data dengan menggunakan metode *random forest classifier* dan *grid search* serta dilakukan *cross validasi* menggunakan *k-fold cross validation* untuk mendapatkan validitas model. Maka, di dapatkan presentase akurasi sebesar 82% pada data dosen dan 85% untuk data tenaga kependidikan.

Hasil output pengklasifikasian adalah akurasi, presisi, TPR (*Recall*), dan *F1-score*. Ditambahkan juga berupa grafik perbandingan dari perbedaan hasil dataset tenaga kependidikan dan dosen. Adapun tabel, gambar dan grafik perbandingan dari hasil klasifikasi, terlihat pada Tabel 1 yang merupakan hasil klasifikasi data tenaga kependidikan, Tabel 2 yang menunjukkan hasil klasifikasi data dosen, Tabel 3 yang merupakan hasil klasifikasi test tenaga kependidikan dan Tabel 4 yang menyajikan hasil klasifikasi data test tenaga dosen. Sedangkan Gambar 2 menyajikan grafik perbandingan hasil *classification report* data tenaga kependidikan (tendik) dan dosen dan Gambar 3 menggambarkan grafik perbandingan hasil klasifikasi test data tenaga kependidikan (tendik) dan dosen

Tabel 1. Hasil *Classification Report* Data Tenaga Kependidikan

	<i>precision</i>	<i>recall</i>	<i>f1-score</i>	<i>support</i>
Tidak Disarankan	0.76	1.00	0.86	16
Dipertimbangkan	0.90	0.95	0.93	20
Disarankan	0.93	0.65	0.76	20
<i>Accuracy</i>			0.86	56
<i>Macro avg</i>	0.87	0.87	0.85	56
<i>Weighted avg</i>	0.87	0.86	0.85	56

Tabel 2. Hasil *Classification Report* Data Dosen

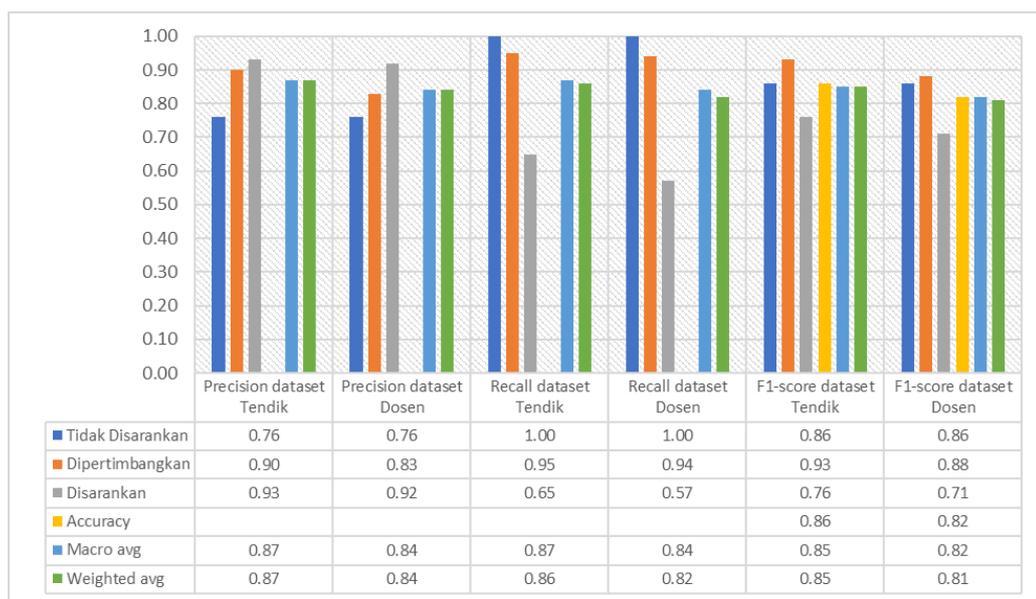
	<i>precision</i>	<i>recall</i>	<i>f1-score</i>	<i>support</i>
Tidak Disarankan	0.76	1.00	0.86	19
Dipertimbangkan	0.83	0.94	0.88	16
Disarankan	0.92	0.57	0.71	21
<i>Accuracy</i>			0.82	56
<i>Macro avg</i>	0.84	0.84	0.82	56
<i>Weighted avg</i>	0.84	0.82	0.81	56

Tabel 3. Hasil Klasifikasi Test Data Tenaga Kependidikan

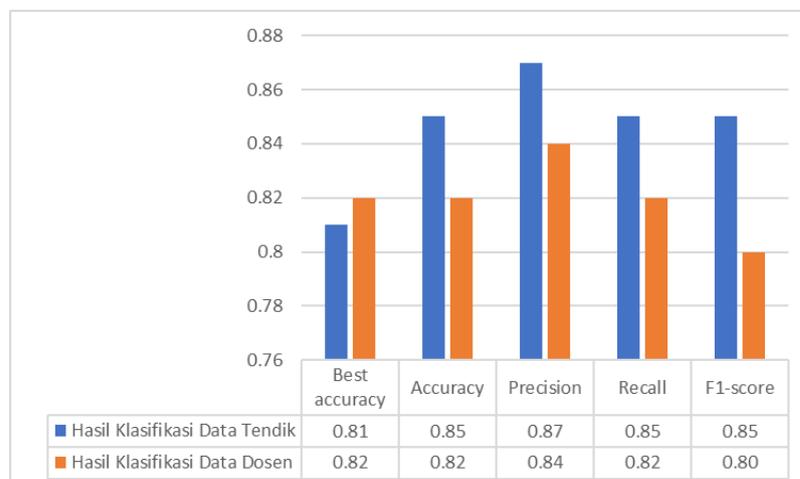
<i>Best accuracy</i>	0.81
<i>Accuracy</i>	0.85
<i>Precision</i>	0.87
<i>Recall</i>	0.85
<i>F1-score</i>	0.85

Tabel 4. Hasil Klasifikasi Test Data Dosen

<i>Best accuracy</i>	0.82
<i>Accuracy</i>	0.82
<i>Precision</i>	0.84
<i>Recall</i>	0.82
<i>F1-score</i>	0.80



Gambar 2. Grafik perbandingan hasil *classification report* data tenaga kependidikan (tendik) dan dosen



Gambar 3. Grafik perbandingan hasil klasifikasi test data tenaga kependidikan (tendik) dan dosen

5. Kesimpulan

Setelah dilakukannya pengolahan data dari hasil psikotes Dosen dan Tenaga kependidikan dengan menggunakan *Random Forest* dan *Grid Search* ditambahkan dengan evaluasi model menggunakan *k-fold cross* validasi dapat diperoleh akurasi yang cukup bagus dengan total akurasi sebesar 85% untuk dataset tenaga kependidikan dan 82% untuk dataset dosen. Dari hasil pengklasifikasian data subjek dosen dan tenaga kependidikan menunjukkan adanya gap atau ketidaksesuaian dikarenakan perbedaan terkait penilaian dari masing-masing aspek antara dosen dan tenaga kependidikan. Kriteria penilaian dosen tidak dapat disamakan dengan tenaga kependidikan.

Penelitian selanjutnya disarankan untuk mengeksplorasi algoritma machine learning lainnya, seperti Gradient Boosting, Support Vector Machines, atau Neural Networks, untuk membandingkan kinerja dan menemukan metode yang paling efektif dalam klasifikasi hasil psikotes. Berikutnya juga penelitian kedepan diharapkan melakukan optimasi hiperparameter yang lebih mendalam dan luas menggunakan teknik yang lebih canggih seperti Bayesian Optimization atau Genetic Algorithms.

6. Daftar Pustaka

- [1] L. Binarwati, I. Mukhlash, and S. Soetrisno, "Implementasi Algoritma Genetika untuk Optimalisasi Random Forest Dalam Proses Klasifikasi Penerimaan Tenaga Kerja Baru: Studi Kasus PT.XYZ," *JURNAL SAINS DAN SENI ITS*, vol. 6, no. 2, pp. A78–A82, 2017.
- [2] N. Zhu, C. Zhu, L. Zhou, Y. Zhu, and X. Zhang, "Optimization of the Random Forest Hyperparameters for Power Industrial Control Systems Intrusion Detection Using an Improved Grid Search Algorithm," *Applied Sciences (Switzerland)*, vol. 12, no. 20, Oct. 2022, doi: 10.3390/app122010456.
- [3] I. Muhamad and M. Matin, "Hyperparameter Tuning menggunakan GridsearchCV pada Random Forest untuk Deteksi Malware," *JURNAL MULTINETICS*, vol. 9, no. 1, 2023.
- [4] D. Genta, I. Desantha, K. Muslim Lhaksana, and D. Richasdy, "Aplikasi Sistem Seleksi Pelamar Kerja dengan menggunakan Metode Random Forest."
- [5] L. Breiman, "Random Forests," *Mach Learn*, vol. 45, pp. 5–32, 2001.
- [6] N. R. Kurnianda, I. Mutia, and I. Hermadi, "Design of Psychological Test Expert System for New Employees Recruitment with Profile Matching Method," European Alliance for Innovation n.o., Feb. 2022. doi: 10.4108/eai.28-10-2020.2315349.
- [7] J. Bergstra, J. B. Ca, and Y. B. Ca, "Random Search for Hyper-Parameter Optimization Yoshua Bengio," 2012. [Online]. Available: <http://scikit-learn.sourceforge.net>.
- [8] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (II)*, Montreal, Quebec, Canada, Aug. 1995, pp. 1137–1143. [Online]. Available: [http://roboticsStanfordedu/"ronnyk](http://roboticsStanfordedu/)
- [9] J. P. Matekenya, "A SWOT Analysis Of Machine Learning in Psychometric Evaluations For Recruitment," 2020.
- [10] H. Jantan, A. R. Hamdan, and Z. A. Othman, "Human Talent Prediction in HRM using C4.5 Classification Algorithm," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 02, no. 08, pp. 2526–2534, 2010, [Online]. Available: <https://www.researchgate.net/publication/50194242>
- [11] P. Refaailzadeh, L. Tang, and H. Liu, "Cross-Validation," in *Encyclopedia of Database Systems*, Boston, MA: Springer US, 2009, pp. 532–538. doi: 10.1007/978-0-387-39940-9_565.
- [12] B. Ghimire, J. Rogan, and J. Miller, "Contextual land-cover classification: Incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic," *Remote Sensing Letters*, vol. 1, no. 1, pp. 45–54, Mar. 2010, doi: 10.1080/01431160903252327.
- [13] A. A. Mahmudi, "Sistem Pendukung Keputusan Rekrutmen Tenaga Kependidikan Menggunakan Metode Promethee," vol. 3, no. 1, pp. 52–62, 2022.
- [14] M. Rianto and R. Yunis, "Analisis Runtun Waktu Untuk Memprediksi Jumlah Mahasiswa Baru Dengan Model Random Forest," *Paradigma - Jurnal Komputer dan Informatika*, vol. 23, no. 1, Mar. 2021, doi: 10.31294/p.v23i1.9781.
- [15] M. Inda Rahayu and K. Ari Sambodo, "Random Forest Classification Of Jambi and South Sumatera Using ALOS PALSAR Data," *International Journal of Remote Sensing and Earth Sciences*, vol. 10, no. 2, pp. 134–141, 2013.