

TUGAS AKHIR
PERBANDINGAN AKURASI ALGORITMA MACHINE
LEARNING UNTUK PREDIKSI PENDAFTAR MAHASISWA
BARU DI UNIVERSITAS XYZ

Diajukan sebagai salah satu syarat untuk memperoleh gelar
Sarjana Komputer di Program Studi Informatika



Oleh :

Muhammad Rohmattullah Joyonegoro

1461600051

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS 17 AGUSTUS 1945 SURABAYA
2020

Halaman ini sengaja dikosongkan

FINAL PROJECT
COMPARASION OF MACHINE LEARNING ALGORITHM
ACCURACY FOR PREDICTION STUDENT ENROLLMENT
AT XYZ UNIVERSITY

Prepared as partial fulfilment of the requirement for the degree of
Sarjana Komputer at Informatics Department



By :

Muhammad Rohmattullah Joyonegoro

1461600051

INFORMATICS DEPARMENT
FACULTY OF ENGINEERING
UNIVERSITAS 17 AGUSTUS 1945 SURABAYA

2020

Halaman ini sengaja dikosongkan

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS 17 AGUSTUS 1945 SURABAYA**

LEMBAR PENGESAHAN TUGAS AKHIR

Nama : Muhammad Rohmattullah Joyonegoro
NBI : 1461600051
Prodi : Informatika
Fakultas : Teknik
Judul : PERBANDINGAN AKURASI ALGORITMA
MACHINE LEARNING UNTUK PREDIKSI
PENDAFTAR MAHASISWA BARU DI
UNIVERSITAS XYZ

Mengetahui / Menyetujui

Dosen Pembimbing



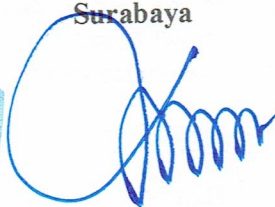
Supangat, M.Kom., ITIL., COBIT
NPP. 20460110602

**Dekan Fakultas Teknik
Universitas 17 Agustus 1945
Surabaya**

**Ketua Program Studi Informatika
Universitas 17 Agustus 1945
Surabaya**



Dr. Ir. Sajiyo M.Kes
NPP. 20410900197



Geri Kusnanto, S.Kom, MM
NPP. 20460940401

Halaman ini sengaja dikosongkan

PERNYATAAN KEASLIAN DAN PERSETUJUAN PUBLIKASI TUGAS

AKHIR

Saya yang bertanda tangan di bawah ini :

Nama : Muhammad Rohmattullah Joyonegoro
NBI : 1461600051
Fakultas / Program Studi : Teknik / Informatika
Judul Tugas Akhir : Perbandingan Algoritma Machine Learning untuk
Prediksi Jumlah Pendaftar Mahasiswa Baru Di
Universitas XYZ

Menyatakan dengan sesungguhnya bahwa :

1. Tugas Akhir dengan judul diatas bukan merupakan tipuan atau duplikasi dari Tugas Akhir yang sudah dipublikasikan dan atau pernah dipakai untuk mendapatkan gelar Sarjana Teknik di lingkungan Universitas 17 Agustus 1945 Surabaya maupun di perguruan tinggi atau instansi manapun. Kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.
2. Tugas Akhir dengan judul diatas bukan merupakan plagiarisme, pencurian hasil karya milik orang lain, hasil kerja orang lain untuk kepentingan saya karena hubungan material maupun non-material, ataupun segala kemungkinan lain yang pada hakekatnya bukan merupakan karya tulis tugas akhir saya secara orisinal dan otentik.
3. Demi pengembangan ilmu pengetahuan saya memberikan hak atas Tugas Akhir ini kepada Universitas 17 Agustus 1945 Surabaya untuk menyimpan mengalihmedia / formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta.
4. Pernyataan ini saya buat dengan kesadaran sendiri dan tidak atas tekanan ataupun paksaan dari pihak manapun demi menegakkan integritas akademik di institusi ini dan bila kemudian hari diduga kuat ada ketidaksesuaian antara fakta dengan kenyataan ini, saya bersedia diproses oleh tim Fakultas yang dibentuk untuk melakukan verifikasi, dengan saksi terberat berupa pembatalan kelulusan / kesarjanaan



Muhammad Rohmattullah Joyonegoro

1461600051

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kepada Allah Yang Maha Esa dan Yang Maha Kuasa yang senantiasa melimpahkan Rahmat dan Hidayah-Nya, sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul “PERBANDINGAN AKURASI ALGORITMA MACHINE LEARNING UNTUK PREDIKSI PENDAFTAR MAHASISWA BARU DI UNIVERSITAS XYZ” sebagai salah satu syarat untuk menyelesaikan studi dan mendapat gelar sarjana di Universitas 17 Agustus 1945 Surabaya. Penulis menyadari bahwa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan hingga sampai penyusunan tugas akhir ini sangatlah penting bagi penulis untuk menyelesaikan dengan baik.

Selain itu penulis ingin menyampaikan ucapan terima kasih kepada pihak – pihak berikut :

1. Ibu tercinta yang selalu memberikan doa, motivasi, serta dukungan kepada penulis.
2. Supangat, M.Kom., ITIL,. COBIT. selaku dosen pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk membantu serta mengarahkan penulis dalam penyusunan tugas akhir ini.
3. TIM Direktorat Sistem Informasi yang telah membantu dalam penyusunan tugas akhir.
4. Gabriele N. Rattu, dan teman – teman lainnya yang telah membantu dan memberi semangat agar penyusunan tugas akhir ini cepat selesai.

Akhir kata, semoga Allah Yang Maha Esa membalas segala kebaikan semua pihak yang telah membantu. Semoga tugas akhir ini membawa manfaat kepada berbagai pihak

Sabtu, 20 Juni 2020

Penulis

Halaman ini sengaja dikosongkan

ABSTRAK

Nama : Muhammad Rohmattullah Joyonegoro
Program Studi : Informatika
Judul : Perbandingan Akurasi Algoritma Machine Learning Untuk
Prediksi Pendaftar Mahasiswa Baru di Universitas XYZ

Penerimaan mahasiswa tahun ajaran baru dapat mengalami peningkatan dan dapat juga mengalami penurunan. Hal ini merupakan suatu masalah yang dihadapi Universitas dalam menentukan langkah-langkah strategi kedepannya. Sehingga diperlukan adanya prediksi atau peramalan untuk mengetahui perolehan jumlah mahasiswa baru, agar semua kebijakan dan keputusan dalam menyusun perencanaan kedepan dapat terpenuhi dengan baik.

Penelitian ini bertujuan untuk mengidentifikasi dan memprediksi jumlah pendaftar mahasiswa baru dengan menggunakan Machine learning. Dataset yang digunakan adalah jumlah pendaftar mahasiswa baru pada tahun 2018-2020 Universitas XYZ dengan jumlah 7485 data yang telah dinormalisasi dengan 6 Data Fitur dan 1 Label Data. Algoritma yang digunakan dalam prediksi ini adalah Gradient Boosting, Decision Tree, K-NN, Logistic Regresion dan Random Forest. Dimana hasil prediksi tersebut dapat memberikan kemudahan kepada Universitas dalam menentukan langkah-langkah strategi dalam mengambil keputusan dan kebijakan pada tahun yang akan datang.

Kata Kunci: *Machine Learning, Performa Accuration Method, Student Enrollment, Supervised Learning.*

Halaman ini sengaja dikosongkan

ABSTRACT

Name : Muhammad Rohmattullah Joyonegoro
Department : Informatics
Title : Comparasion Accuracy Of Machine Learning Algorithm For
Prediction Student Enrollment At XYZ University

Admission of New school year students can be improved and can also be decreased. This is an issue facing the university in determining the future strategy steps. So there is a need for predictions or forecasting to determine the acquisition of new students, so that all policies and decisions in making future planning can be fulfilled properly.

This research aims to identify and predict the number of new student registries using Machine learning. The Dataset used was the number of new student registries in the year 2018-2020 XYZ University with a total of 7485 data that had been normalized with 6 feature Data and 1 Data Label. The algorithms used in this predisksi are Gradient Boosting, Decision Tree, K-NN, Logistic Regression and Random Forest. Where the prediction results can provide ease to the university in determining the strategy measures in making decisions and policies in the coming year.

Kata Kunci: *Machine Learning, Performa Accuration Method, Student Enrollment, Supervised Learning.*

DAFTAR ISI

KATA PENGANTAR	ix
ABSTRAK	xi
DAFTAR ISI	xiv
DAFTAR TABEL	xvii
DAFTAR GAMBAR	xviii
DAFTAR LAMPIRAN	xix
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah	2
1.4. Tujuan Penelitian	2
BAB 2 KAJIAN PUSTAKA DAN DASAR TEORI	3
2.1. Kajian Pustaka	3
2.2. Machine Learning	4
2.3. Gradient Boosting	5
2.4. Decision Tree	6
2.5. K-NN	8
2.6. Logistic Regression	8
2.7. Random Forest	9
2.8. Python	11
2.9. Scikit-Learn	11
BAB 3 METODOLOGI PENELITIAN	13
3.1. Perancangan Sistem	13
3.1.1. Cara Kerja Machine Learning	13
3.1.2. Teknik Machine Learning	14
3.1.3. Cara Kerja Supervised Learning	15
3.1.4. Flowchart Supervised Learning	16
3.1.5. Flowchart Machine Learning Perbandingan Tingkat Akurasi Prediksi	17

3.1.6.	Flowchart Metode K-NN	18
3.1.7.	Flowchart Metode Decision Tree.....	19
3.1.8.	Flowchart Metode Random Forest.....	20
3.1.9.	Flowchart Metode Logistic Lienar Regresion.....	20
3.1.10.	Flowchart Metode Gradient Boosting	21
3.1.11.	Flowchart Metode Prediksi	22
3.2.	Implementasi	23
3.2.1.	Tools.....	23
3.2.2.	Lingkungan Uji Coba.....	23
BAB 4 HASIL DAN PEMBAHASAN.....		25
4.1.	Ujicoba Sistem	25
4.1.1.	Data Penerimaan Mahasiswa Baru.....	25
4.1.2.	Alur Proses Import Dataset dan Overview Data	26
4.1.3.	Hasil Ujicoba Dengan Gradient Boosting	30
4.1.4.	Hasil Ujicoba Dengan Decision-Tree	31
4.1.5.	Hasil Ujicoba Dengan K-NN	31
4.1.6.	Hasil Ujicoba Dengan Logistic-Regresion.....	32
4.1.7.	Hasil Ujicoba Dengan Random Forest.....	32
4.2.	Hasil Akurasi Tiap Metode.....	33
4.3.	Cross Validation dan Perhitungan Akurasi dan Missclasification.....	33
4.3.1.	Import Librari dan Read Data	33
4.3.2.	Splitting Dataset.....	34
4.3.3.	Scaling Feature.....	35
4.3.4.	Parameter Evaluation	36
4.3.5.	Adjusting Development Thereshold	40
4.3.6.	Prediction Test Set Result	42
4.3.7.	Making Confusion Matrix.....	44
4.3.8.	Ploting Prediction and Result.....	50
4.4.	Hasil Dari Uji Coba.....	60
4.4.1.	Result Uji Coba Akurasi	60
BAB V PENUTUP.....		68

5.1. Kesimpulan.....	68
5.2. Saran.....	68
DAFTAR PUSTAKA	70
DAFTAR LAMPIRAN	71

DAFTAR TABEL

Tabel 3.2.1 Persiapan Tols yang digunakan	23
Tabel 4.2.1 Hasil Akurasi Prediksi Tiap Metode	33
Tabel 4.4.1 Hasil Uji Coba 100 Data	60
Tabel 4.4.2 Hasil Uji Coba 500 Data.....	61
Tabel 4.4.3 Hasil Uji Coba 1500 Data.....	62
Tabel 4.4.4 Hasil Uji Coba 3000 Data.....	62
Tabel 4.4.5 Hasil Uji Coba 5000 Data.....	63
Tabel 4.4.6 Hasil Uji Coba Dataset	64

DAFTAR GAMBAR

Gambar 3.1.1 Cara Kerja Machine Learning	13
Gambar 3.1.2 Metode Pemisah Data Supervised dan Unsupervised	14
Gambar 3.1.3 Cara Kerja Supervised Training	15
Gambar 3.1.4 Flowchart Supervised Learning	16
Gambar 3.1.5 Flowchart Alur Sistem	17
Gambar 3.1.6 Flowchart Metode K-NN	18
Gambar 3.1.7 Flowchart Metode Decision Tree	19
Gambar 3.1.8 Flowchart Metode Random Forest	20
Gambar 3.1.9 Flowchart General Data Metode Linear Regression	20
Gambar 3.1. 10 Flowcart Metode Gradient Boosting	21
Gambar 3.1.11 Flowchart Prediksi Jumlah Pendaftar Mahasiswa Baru	22
Gambar 4.1.1 Data Penerimaan Mahasiswa Baru	25
Gambar 4.1.2 Dataset yang digunakan	25
Gambar 4.1.3 Keterangan Dataset	26
Gambar 4.1.4 Source untuk load Library	26
Gambar 4.1.5 Source load Dataset	27
Gambar 4.1.6 Source Informasi Dataset	27
Gambar 4.1.7 Source Dimensi data dan Grup by data pendaftar	27
Gambar 4.1.8 Penjelasan Basic EDA dari data fitur label pada dataset	28
Gambar 4.1.9 Image Pairplot Data Fitur	29
Gambar 4.1.10 Analisis Data menggunakan Diagram	29
Gambar 4.1.11 Split Dataset Uji Coba Pertama	30
Gambar 4.1.12 Hasil Prediksi Gradient Boosting	30
Gambar 4.1.13 Hasil Prediksi Menggunakan Metode Decision Tree	31
Gambar 4.1.14 Hasil Prediksi Metode K-NN	31
Gambar 4.1.15 Hasil Prediksi Menggunakan Metode Logistic-Regresion ...	32
Gambar 4.1. 16 Hasil Prediksi Menggunakan Metode Random Forest	32
Gambar 4.4.1 Grafik Perbandingan Akurasi 100 Data	61
Gambar 4.4.2 Grafik Perbandingan Akurasi 500 Data	61
Gambar 4.4.3 Grafik Perbandingan Akurasi 1500 Data	62
Gambar 4.4.4 Grafik Perbandingan Akurasi 3000 Data	63
Gambar 4.4.5 Grafik Perbandingan Akurasi 5000 Data	63
Gambar 4.4.6 Grafik Kesimpulan Perbandingan Akurasi 5 Metode	64

DAFTAR LAMPIRAN

Lampiran 1 : Dokumentasi Proses Evaluasi di Google Colab	71
Lampiran 2 : Dokumentasi Akurasi Metode	72

