

DAFTAR ISI

HALAMAN PENGESAHAN.....	ii
PERNYATAAN BEBAS PLAGIASI	iii
HALAMAN PERSEMBAHAN.....	iv
KATA PENGANTAR.....	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR SINGKATAN	xii
INTISARI.....	xiii
ABSTRACT	xiv
BAB I. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Tugas Akhir	3
1.4 Tujuan Tugas Akhir	3
1.5 Manfaat Tugas Akhir	4
1.6 Sistematika Penulisan	4
BAB II. TINJAUAN PUSTAKA DAN DASAR TEORI	6
2.1 Tinjauan Pustaka	6
2.2 Landasan Teori	7
2.2.1 Data	7
2.2.2 Data Science	7
2.2.3 Kecerdasan Buatan	8
2.2.4 Machine Learning	9
2.2.5 Support Vector Machine (SVM)	9
2.2.6 Decision Tree	12
2.2.7 Random Forest	13
2.2.8 Logistic Regression	14
2.2.9 Naive Bayes	15
2.2.10 K-Nearest Neighbors	16
2.2.11 Cross Validation	16
2.2.12 Python	18
2.2.13 Confusion Matrix	19

2.2.14	<i>Scikit-learn</i>	19
2.2.15	<i>Imbalanced-learn</i>	20
2.2.16	Numpy	21
2.2.17	Pandas.....	21
2.2.18	Seaborn.....	21
BAB III. METODE TUGAS AKHIR		23
3.1	Alat dan Bahan Tugas akhir	23
3.1.1	Alat Tugas akhir	23
3.1.2	Bahan Tugas Akhir	24
3.2	Alur Tugas Akhir	25
3.3	Pengumpulan Data	26
3.4	Memahami Data	27
3.5	<i>Cleaning Data</i>	27
3.6	<i>Preparing Data</i>	28
3.7	Merancang Model.....	28
3.8	Evaluasi Model	28
BAB IV. CLASSIFICATION MODEL		29
4.1	<i>Preprocessing Data</i>	29
4.1.1	Mengubah Bentuk Data <i>Moodle Log</i>	30
4.1.2	Mengubah Nama Mahasiswa Menjadi Huruf Kecil	31
4.1.3	<i>Joining Data Moodle Log, Log Netacad, dan Nilai</i>	32
4.1.4	<i>Cleaning Data Dari Noise Dan Outlier</i>	33
4.1.5	Standarisasi Nama Fitur.....	35
4.2	Pemilihan Fitur.....	36
4.3	Pembuatan Target.....	38
4.4	Pembagian <i>Data Training dan Data Test</i>	38
4.5	<i>Oversampling Data</i>	38
4.6	Pembuatan Model.....	40
4.6.1	<i>Support Vector Machine</i>	40
4.6.2	Logistic Regression.....	41
4.6.3	<i>Random Forest</i>	42
4.6.4	<i>K Nearest Neighbors</i>	42
4.6.5	<i>Naive Bayes</i>	44
4.6.6	<i>Decision Tree</i>	46
4.7	Pemilihan Model Terbaik.....	46
BAB V. KESIMPULAN DAN SARAN		49



Learning Analytics: Implementasi Dan Visualisasi Machine Learning Untuk Memprediksi Hasil Belajar Pada Learning Management System (LMS) - Classification Model

AHMAD IHSAN, Dr. Sri Suning Kusumawardani, S.T. M.T.; Ir. Markus Nurtiantara Aji, M.T.

Universitas Gadjah Mada, 2020 | Diunduh dari <http://etd.repository.ugm.ac.id/>

UNIVERSITAS
GADJAH MADA

3.1	Kesimpulan	49
5.2	Saran.....	49
DAFTAR PUSTAKA		51
LAMPIRAN.....		55



DAFTAR TABEL

Tabel 2.1 Hubungan antara True Values dan Prediction	19
Tabel 4.1 Standarisasi Nama Fitur	35
Tabel 4.2 Confusion Matrix	46
Tabel 4.3 Evaluasi Akurasi	47

DAFTAR GAMBAR

Gambar 2.1 Pembuatan Hyperplane Optimal [26].....	10
Gambar 2.2 Perbandingan antara margin kecil (a) dan margin besar (b) dalam pembuatan hyperplane yang optimal [26]	11
Gambar 2.3 Algoritme Random Forest [31]	13
Gambar 2.4 Proses evaluasi cross validation dengan six-folds.....	18
Gambar 3.1 Flowchart Alur Tugas Akhir	25
Gambar 4.1 Data Log Aktivitas eLOK	29
Gambar 4.2 Data Log Netacad.....	30
Gambar 4.3 Data Nilai Akhir Mahasiswa	30
Gambar 4.4 Mengubah Data Menggunakan Tableau.....	31
Gambar 4.5 Source Code untuk Mengubah Nama Mahasiswa menjadi Huruf Kecil.....	31
Gambar 4.6 Nama Sebelum Diubah.....	32
Gambar 4.7 Nama Setelah Diubah	32
Gambar 4.8 Inner Join Menggunakan Tableau	33
Gambar 4.9 Inner Join Menggunakan Tableau (2)	33
Gambar 4.10 Peserta dengan Data Tidak Normal.....	34
Gambar 4.11 Peserta dengan Dua Nilai Berbeda	34
Gambar 4.12 Peserta dengan Nilai Melebihi Batas Maksimal.....	34
Gambar 4.13 Bentuk Akhir Data.....	35
Gambar 4.14 Kolom-kolom yang Dipilih untuk Diseleksi menjadi Fitur.....	36
Gambar 4.15 Heatmap Korelasi Antar Fitur	37
Gambar 4.16 Source Code untuk Pembagian Kategori.....	38
Gambar 4.17 Train Test Split Data	38
Gambar 4.18 Data Imbalanced.....	39



Gambar 4.19 Source Code ADASYN	40
Gambar 4.20 Parameter Terbaik berdasarkan GridsearchCV	40
Gambar 4.21 Evaluasi SVM.....	41
Gambar 4.22 Parameter Terbaik Berdasarkan GridsearchCV	41
Gambar 4.23 Hasil Evaluasi Model Logistic Regression	41
Gambar 4.24 Parameter Terbaik Berdasarkan GridsearchCV	42
Gambar 4.25 Hasil Evaluasi Model Random Forest.....	42
Gambar 4.26 Parameter Terbaik Berdasarkan GridsearchCV	43
Gambar 4.27 Evaluasi Model KNN	43
Gambar 4.28 Grafik Pengaruh n_neighbors pada Akurasi	43
Gambar 4.29 Evaluasi Bernoulli NB.....	44
Gambar 4.30 Evaluasi Multinomial NB.....	44
Gambar 4.31 Evaluasi Complement NB	45
Gambar 4.32 Evaluasi Gaussian NB	45
Gambar 4.33 Parameter Terbaik Decision Tree.....	46
Gambar 4.34 Evaluasi Decision Tree.....	46
Lampiran 1 Summary Data	55
Lampiran 2 Source Code SVM	55
Lampiran 3 Source Code Random Forest	55
Lampiran 4 Source Code KNN	56
Lampiran 5 Source Code Naïve Bayes	56
Lampiran 6 Source Code Decision Tree	57