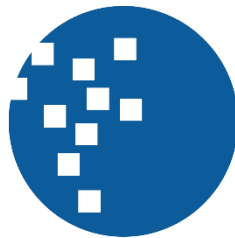


**KOMPARASI OPTIMASI MENGGUNAKAN *LATENT
SEMANTIC ANALYSIS* DAN *LATENT DIRICHLET
ALLOCATION* PADA ANALISIS SENTIMEN BERBASIS
ASPEK DARI ULASAN *ONLINE FOOD DELIVERY (OFD)*
DENGAN SVM**



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Wiky Hendra

00000043861

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2024

KOMPARASI OPTIMASI MENGGUNAKAN *LATENT SEMANTIC ANALYSIS* DAN *LATENT DIRICHLET ALLOCATION* PADA ANALISIS SENTIMEN BERBASIS ASPEK DARI ULASAN *ONLINE FOOD DELIVERY (OFD)* DENGAN SVM



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer (S. Kom.)

Wiky Hendra

00000043861

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNIK DAN INFORMATIKA

UNIVERSITAS MULTIMEDIA NUSANTARA

TANGERANG

2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Wiky Hendra

Nomor Induk Mahasiswa : 00000043861

Program studi : Sistem Informasi

Proposal skripsi dengan judul:

“Komparasi Optimasi Menggunakan *Latent Semantic Analysis* dan *Latent Dirichlet Allocation* Pada Analisis Sentimen Berbasis Aspek Dari Ulasan *Online Food Delivery* (OFD) Dengan SVM”

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi mfaupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas Akhir yang telah saya tempuh.

Tangerang, 28 Mei 2024



Wiky Hendra

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

HALAMAN PENGESAHAN

Skripsi dengan judul

"Komparasi Optimasi Menggunakan *Latent Semantic Analysis* dan *Latent Dirichlet Allocation* Pada Analisis Sentimen Berbasis Aspek Dari Ulasan *Online Food Delivery* (OFD) Dengan SVM"

Oleh

Nama : Wiky Hendra
NIM : 00000043861
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika


Telah diujikan pada hari 28 Mei 2024
Pukul 13.00 s.d 15.00 dan dinyatakan
LULUS

Dengan susunan penguji sebagai berikut:

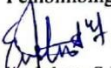
Ketua Sidang


Johan Setiawan, S. Kom., MM.
327106402

Penguji


Dr. Santo Fernandi Wijaya, S. Kom., M.M.
0310016902

Pembimbing


Monika Evehin Johan, S.Kom., M.M.S.I.
0327059501

Ketua Program Studi Sistem Informasi


Ririn Ikana Desanti, S.Kom., M.Kom.
313058001

iii

Komparasi Optimasi Menggunakan..., Universitas Multimedia Nusantara

UNIVERSITAS
MULTIMEDIA
NUSANTARA

LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHASISWA

Yang bertanda tangan dibawah ini:

Nama : Wiky Hendra
Nomor Induk Mahasiswa : 00000043861
Program Studi : Sistem Informasi
Jenjang : S1
Judul Karya Ilmiah : “Komparasi Optimasi Menggunakan *Latent Semantic Analysis* dan *Latent Dirichlet Allocation* Pada Analisis Sentimen Berbasis Aspek Dari Ulasan *Online Food Delivery* (OFD) Dengan SVM”

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial. Saya tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.

Tangerang, 28 Mei 2024



(Wiky Hendra)

* Jika tidak bisa membuktikan LoA jurnal/HKI selama 6 bulan kedepan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

KATA PENGANTAR

Puji Syukur penulis haturkan atas selesainya penulisan Skripsi ini dengan judul “Komparasi Optimasi Menggunakan *Latent Semantic Analysis* Dan *Latent Dirichlet Allocation* Pada Analisis Sentimen Berbasis Aspek Dari *Online Food Delivery* (OFD) Reviews Dengan SVM”. Penulisan ini dilakukan untuk memenuhi salah satu syarat untuk mencapai Gelar Strata 1 Jurusan Sistem Informasi pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya memahami bahwa penulisan ini tidak bisa dilakukan sendiri, sehingga memerlukan berbagai pihak untuk membantu peneliti menyelesaikan penulisan Skripsi. Oleh karena itu, penulis mengucapkan terima kasih kepada :

1. Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Johan Setiawan, S.Kom., M.M., M.B.A, sebagai dosen MRSI yang telah memberikan bimbingan dan arahan dalam penyusunan proposal skripsi ini.
5. Monika Evelin Johan, S.Kom., M.M.S.I., sebagai pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya skripsi ini.
6. Keluarga dan teman-teman saya yang telah memberikan bantuan dukungan informasi, material, dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga karya ilmiah ini dapat bermanfaat bagi pembaca sebagai sumber informasi ataupun inspirasi.

Tangerang, 15 Mei 2024



Wiky Hendra

UNIVERSITAS
MULTIMEDIA
NUSANTARA

**KOMPARASI OPTIMASI MENGGUNAKAN LATENT
SEMANTIC ANALYSIS DAN LATENT DIRICHLET
ALLOCATION PADA ANALISIS SENTIMEN BERBASIS
ASPEK DARI ONLINE FOOD DELIVERY (OFD) REVIEWS
DENGAN SVM**

Wiky Hendra

ABSTRAK

Bagi para pebisnis khususnya di industri *Food and Beverages*, analisis sentimen memberikan hasil yang sangat baik bagi perusahaannya. Dengan industri FnB yang berkembang, pastinya akan memberikan *value* yang terbaik khususnya pada pelayanan. Dengan dengan era yang digital, restoran pun memberikan pelayanan *order* secara *online* disebut *Online Food Delivery* (OFD). Perlu pemahaman opini pelanggan berdasarkan pelayanan yang telah didapat. Salah satu cara untuk memahami opini yaitu analisis sentimen. Namun, analisis sentimen masih belum cukup memberikan pengetahuan hal spesifik apa yang menonjol dan kurang. Perlunya aspek yang dijual oleh restoran, yaitu 4P (*Product, Price, Place, Prmotion*) *Marketing Mix. Aspect-Based Sentiment Analysis* akan dilakukan kepada ulasan pelanggan salah satu restoran dengan (OFD). Proses pekerjaan akan berdasarkan framework CRISP-DM dengan dukungan pemahaman tentang *value* berupa 4P yang dijual oleh restoran OFD. Dilakukannya *scraping* pada *website* dan melakukan pemrosesan pada hasil *scraping* dengan melakukan *text pre-processing*. Selanjutnya, dilakukan pembangunan model dengan membandingkan algoritma *Support Vector Machine* yang ditambahkan *Latent Semantic Analysis* dan *Latent Dirichlet Allocation*. Data diambil dari bulan April 2024 dengan jumlah data sebanyak 2820 ulasan. Berdasarkan model yang telah dibuat, LDA lebih baik daripada LSA dengan selisih 3 persen. Hasil berupa aplikasi *web service* sebuah sentimen untuk pemilik outlet sebagai opini pengambilan keputusan untuk memajukan *value* berdasarkan 4 aspek.

Kata kunci: *Aspect-Based Sentiment Analysis, Latent Semantic Analysis, Latent Dirichlet Allocation, Online Food Delivery, Support Vector Machine.*

**COMPARATIVE OPTIMIZATION USING LATENT SEMANTIC
ANALYSIS AND LATENT DIRICHLET ALLOCATION ON
ASPECT-BASED SENTIMENT ANALYSIS OF ONLINE FOOD
DELIVERY (OFD) REVIEWS
WITH SVM**

Wiky Hendra

ABSTRACT (English)

For business people, especially in the Food and Beverages industry, sentiment analysis provides excellent results for the company. With the growing FnB industry, it will certainly provide the best value, especially in service. With the digital era, restaurants also provide online order services called Online Food Delivery (OFD). It is necessary to understand customer opinions based on the services that have been obtained. One way to understand opinions is sentiment analysis. However, sentiment analysis is still not enough to provide knowledge of what specific things stand out and are lacking. The need for aspects sold by restaurants, namely the 4P (Product, Price, Place, Promotion) Marketing Mix. Aspect-Based Sentiment Analysis will be performed on customer reviews of one restaurant with (OFD). The work process will be based on the CRISP-DM framework with the support of an understanding of the value in the form of 4Ps sold by OFD restaurants. Scraping is done on the website and processing the scraping results by doing text pre-processing. Furthermore, model building is carried out by comparing the Support Vector Machine algorithm added with Latent Semantic Analysis and Latent Dirichlet Allocation. Data was taken from April 2024 with a total of 2820 reviews. Based on the model that has been created, LDA is better than LSA with a difference of 3 percent. The result is a sentiment web service application for outlet owners as a decision-making opinion to advance value based on 4 aspects.

Keywords: *Aspect-Based Sentiment Analysis, Latent Semantic Analysis, Latent Dirichlet Allocation, Online Food Delivery, Support Vector Machine.*

M U L T I M E D I A
N U S A N T A R A

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT.....	ii
HALAMAN PENGESAHAN.....	iii
LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHASISWA	iv
KATA PENGANTAR	v
ABSTRAK.....	vi
ABSTRACT (English).....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR RUMUS	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN.....	xvii
jabab I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan dan Manfaat Penelitian	6
1.4.1 Tujuan Penelitian	6
1.4.2 Manfaat Penelitian	6
1.5 Sistematika Penulisan	7
BAB II LANDASAN TEORI.....	8
2.1 Penelitian Terdahulu	8
2.2 Objek Penelitian.....	16
2.2.1 Food Service	16
2.2.1.1 Restoran	18
2.2.1.1.1 Fast Food Restaurant.....	19
2.2.2 Marketing Mix	20

2.2.2.1	Product	20
2.2.2.2	Price.....	21
2.2.2.3	Place	21
2.2.2.4	Promotion.....	21
2.2.3	Sentiment Analysis	21
2.2.4	Aspect-Based Sentiment Analysis (ABSA).....	23
2.3	Algoritma / Framework.....	23
2.3.1	CRISP-DM.....	23
2.3.2	Latent Semantic Analysis (LSA)	27
2.3.3	Latent Dirichlet Allocation (LDA).....	28
2.3.4	Support Vector Machine (SVM).....	29
2.4	Tools.....	30
2.4.1	Google Colab	30
BAB III METODOLOGI PENELITIAN		31
3.1	Gambaran Umum Objek Penelitian	31
3.1.1	Outlet <i>Online Food Delivery</i> (OFD)	31
3.2	Alur Penelitian	33
3.3	Metode Penelitian.....	37
3.4	Teknik Pengumpulan Data.....	38
3.4.1	Data Collection.....	38
3.5	Teknik Analisis Data.....	41
BAB IV ANALISIS DAN HASIL PENELITIAN		44
4.1	Analisa Masalah dan Kebutuhan Penelitian.....	44
4.1.1	Tahap Business Understanding	44
4.1.2	Tahap Data Understanding.....	45
4.1.3	Tahap Data Preparation.....	51
4.1.4	Tahap Modelling	60
4.1.4.1.	Model Latent Semantic Analysis (LSA)	60
4.1.4.2.	Model Latent Dirichlet Allocation (LDA).....	61
4.1.4.3.	K-Fold Cross Validation	62
4.1.5	Tahap Evaluation	63
4.1.6	Tahap Deployment	71

4.2 Hasil Analisis Data dan Diskusi.....	73
BAB V SIMPULAN DAN SARAN.....	77
5.1 Simpulan.....	77
5.2 Saran.....	77
DAFTAR PUSTAKA.....	80
LAMPIRAN.....	84



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR TABEL

Tabel 2. 1 Tabel berisi artikel jurnal penelitian terdahulu	8
Tabel 2. 2 Tipe dari operasi Food Service	16
Table 3. 1 Perbandingan Model SVM, Naïve Bayes, dan LR.....	38
Table 3. 2 Perbandingan antara CRISP-DM dan SEMMA	41
Table 3. 3 Perbandingan antara Google Colab dan Jupyter Notebook	42
Table 4. 1 Penjelasan variabel yang digunakan pada dataset.....	50
Table 4. 2 Hasil proses dari 4 aspek dengan hasil kernel, akurasi normal, dan akurasi K-Fold.....	75

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR RUMUS

Rumus 2. 1 Rumus TF-IDF.....	26
Rumus 2. 2 Rumus Singular Value Decomposition (SVD).....	27
Rumus 2. 3 Ekuasi form Klasifikasi	29
Rumus 2. 4 Rumus hyperplane	29



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 1. 1 Nilai dan Pertumbuhan PDB Industri Makanan dan Minuman (2010-2021)	1
Gambar 2. 1 Gambaran keseluruhan mengenai Sentimen Analisis.....	22
Gambar 2. 2 Model Grafik untuk Latent Dirichlet Allocation.....	28
Gambar 2. 3 Gambaran Support Vector Machine.....	29
Gambar 3. 1 Flowchart alur penelitian secara keseluruhan berdasarkan framework CRISP-DM.....	33
Gambar 3. 2 Hasil penggunaan library BeautifulSoup dan requests untuk automatic scraping dengan hasil Response 403 berarti Forbidden	35
Gambar 3. 3 Potongan flowchart untuk proses Data Preparation berdasarkan Gambar 3. 1	36
Gambar 3. 4 Pembagian proses pekerjaan berdasarkan Flowchart Data Collection.	39
Gambar 3. 5 Flowchart proses detil pengolahan API menjadi data yang siap digunakan.	40
Gambar 4. 1 Page salah satu outlet pada platform GoFood.....	44
Gambar 4. 2 API yang memberikan seluruh data menuju Front-End.....	46
Gambar 4. 3 Gambar proses pengambilan API dengan copy cURL as bash.....	46
Gambar 4. 4 Halaman Page Postman agar siap ke proses selanjutnya	47
Gambar 4. 5 Tampilan page ketika klik Import. Page meminta untuk memasukkan cURL dengan apa yang kita copy dari Inspect Element.	47
Gambar 4. 6 API Sudah terlihat. Berikut adalah API GoFood untuk mengambil review dari McDonald's SDC.....	48
Gambar 4. 7 Export API yang telah disave dengan format Collection v2.1. File akan terbentuk sebagai .json format.....	48
Gambar 4. 8 Import Postman pada Katalon Studio	49
Gambar 4. 9 Gambar potongan code untuk Scraping di Katalon Studio	49
Gambar 4. 10 Hasil data scraping dalam bentuk excel	51
Gambar 4. 11 Instalasi yang dibutuhkan code untuk melakukan translasi teks... ..	52

Gambar 4. 12 Potongan Code mengenai translasi teks dan menyimpan hasil data pada excel.....	52
Gambar 4. 13 Potongan Code mengenai translasi teks dan menyimpan hasil data pada excel.....	52
Gambar 4. 14 Potongan code inisiasi menggunakan excel yang sudah ditranslasi dan drop NA pada Tags.....	53
Gambar 4. 15 Potongan Code mengenai labeling otomatis menggunakan TextBlob dan mengukur menggunakan polarity.....	53
Gambar 4. 16 Semua tagging yang ada beserta total tagging.....	54
Gambar 4. 17 Beberapa Tags yang mewakili aspek-aspek yang telah disediakan.	55
Gambar 4. 18 Total data yang akan diolah setiap aspek.....	55
Gambar 4. 19 Potongan Code untuk Case Folding pertama.....	56
Gambar 4. 20 Potongan Code untuk Case Folding kedua.....	56
Gambar 4. 21 Hasil distribusi labeling pada Aspect Product.....	59
Gambar 4. 22 Hasil distribusi labeling pada Aspect Price.....	59
Gambar 4. 23 Hasil distribusi labeling pada Aspect Place.....	59
Gambar 4. 24 Potongan Code mengenai vektorisasi menggunakan TF-IDF dengan optimasi n-gram.....	59
Gambar 4. 25 Potongan code pemodelan LSA.....	60
Gambar 4. 26 Potongan code pemodelan LDA.....	61
Gambar 4. 27 Potongan code untuk melakukan K-Fold pada model LSA dan LDA.....	62
Gambar 4. 28 Classification report pada LSA beserta parameter kernel yang terbaik pada aspek Produk.....	63
Gambar 4. 29 Classification report pada LDA beserta parameter kernel yang terbaik pada aspek Produk.....	63
Gambar 4. 30 Hasil rata-rata akurasi, presisi, recall, dan f1-score pada k-fold cross validation menggunakan parameter yang sama untuk Latent Semantic Analysis	64

Gambar 4. 31 Hasil rata-rata akurasi, presisi, recall, dan f1-score pada k-fold cross validation menggunakan parameter yang sama untuk Latent Dirichlet Allocation.....	64
Gambar 4. 32 Line Chart mengenai rata-rata akurasi setiap iterasi K-Fold Cross Validation, dimulai dari K = 2 hingga K = 29 pada Product	64
Gambar 4. 33 Classification report pada LSA beserta parameter kernel yang terbaik pada aspek Price.....	65
Gambar 4. 34 Classification report pada LDA beserta parameter kernel yang terbaik pada aspek Price.....	65
Gambar 4. 35 Hasil rata-rata akurasi, presisi, recall, dan f1-score pada k-fold cross validation menggunakan parameter yang sama untuk Latent Semantic Analysis pada aspek Price.....	65
Gambar 4. 36 Hasil rata-rata akurasi, presisi, recall, dan f1-score pada k-fold cross validation menggunakan parameter yang sama untuk Latent Dirichlet Allocation pada aspek Price	66
Gambar 4. 37 Line Chart mengenai rata-rata akurasi setiap iterasi K-Fold Cross Validation, dimulai dari K = 2 hingga K = 29 pada Price	66
Gambar 4. 38 Classification report pada LSA dan LDA beserta parameter kernel yang terbaik pada aspek Promotion	67
Gambar 4. 39 Hasil rata-rata akurasi, presisi, recall, dan f1-score pada k-fold cross validation menggunakan parameter yang sama untuk Latent Semantic Analysis pada aspek Promotion	67
Gambar 4. 40 Hasil rata-rata akurasi, presisi, recall, dan f1-score pada k-fold cross validation menggunakan parameter yang sama untuk Latent Dirichlet Allocation pada aspek Promotion	67
Gambar 4. 41 Line Chart mengenai rata-rata akurasi setiap iterasi K-Fold Cross Validation, dimulai dari K = 2 hingga K = 29 pada Promotion.....	68
Gambar 4. 42 Classification report pada LSA beserta parameter kernel yang terbaik pada aspek Place	69
Gambar 4. 43 Classification report pada LDA beserta parameter kernel yang terbaik pada aspek Place	69

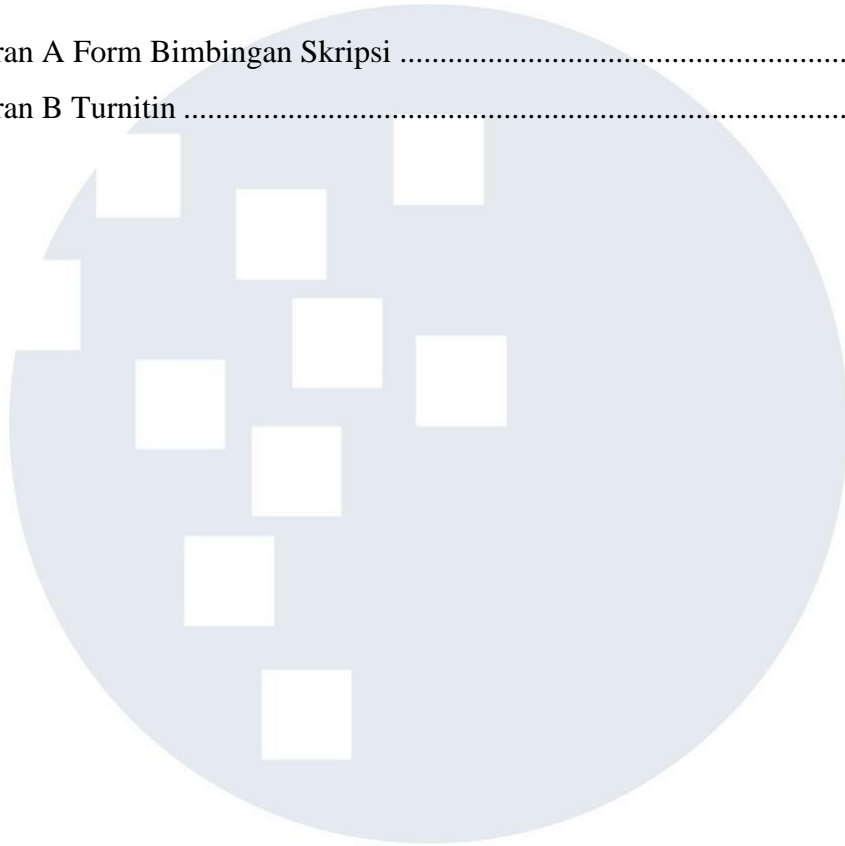
Gambar 4. 44 Hasil rata-rata akurasi, presisi, recall, dan f1-score pada k-fold cross validation menggunakan parameter yang sama untuk Latent Semantic Analysis pada aspek Place	69
Gambar 4. 45 Hasil rata-rata akurasi, presisi, recall, dan f1-score pada k-fold cross validation menggunakan parameter yang sama untuk Latent Dirichlet Allocation pada aspek Place.....	69
Gambar 4. 46 Line Chart mengenai rata-rata akurasi setiap iterasi K-Fold Cross Validation, dimulai dari K = 2 hingga K = 29 pada Place	70
Gambar 4. 47 Potongan code untuk melakukan export model pada file format .pkl,.....	71
Gambar 4. 48 Potongan code untuk melakukan export model pada file format .pkl,.....	72

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran A Form Bimbingan Skripsi	84
Lampiran B Turnitin	85



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA