

## DAFTAR ISI

HALAMAN SAMBUNG .....	i
HALAMAN JUDUL.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERNYATAAN KEASLIAN .....	v
HALAMAN PERSEMBAHAN .....	vi
HALAMAN MOTTO .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiv
DAFTAR LAMPIRAN.....	xv
DAFTAR ISTILAH .....	xvi
INTISARI.....	xvii
ABSTRACT.....	xviii
<b>BAB I PENDAHULUAN</b>	
A. Latar Belakang Masalah .....	1
B. Rumusan Masalah.....	5
C. Batasan Masalah .....	5
D. Tujuan Penelitian .....	6
E. Manfaat Penelitian .....	6
<b>BAB II TINJAUAN PUSTAKA</b>	
A. Landasan Teori.....	7
B. Penelitian Sebelumnya.....	14
<b>BAB III METODE PENELITIAN</b>	
A. Tempat dan Waktu Penelitian.....	17
B. Metode Pengumpulan Data.....	17
C. Alat dan Bahan Penelitian.....	18
D. Konsep Penelitian .....	20

## BAB IV HASIL DAN PEMBAHASAN

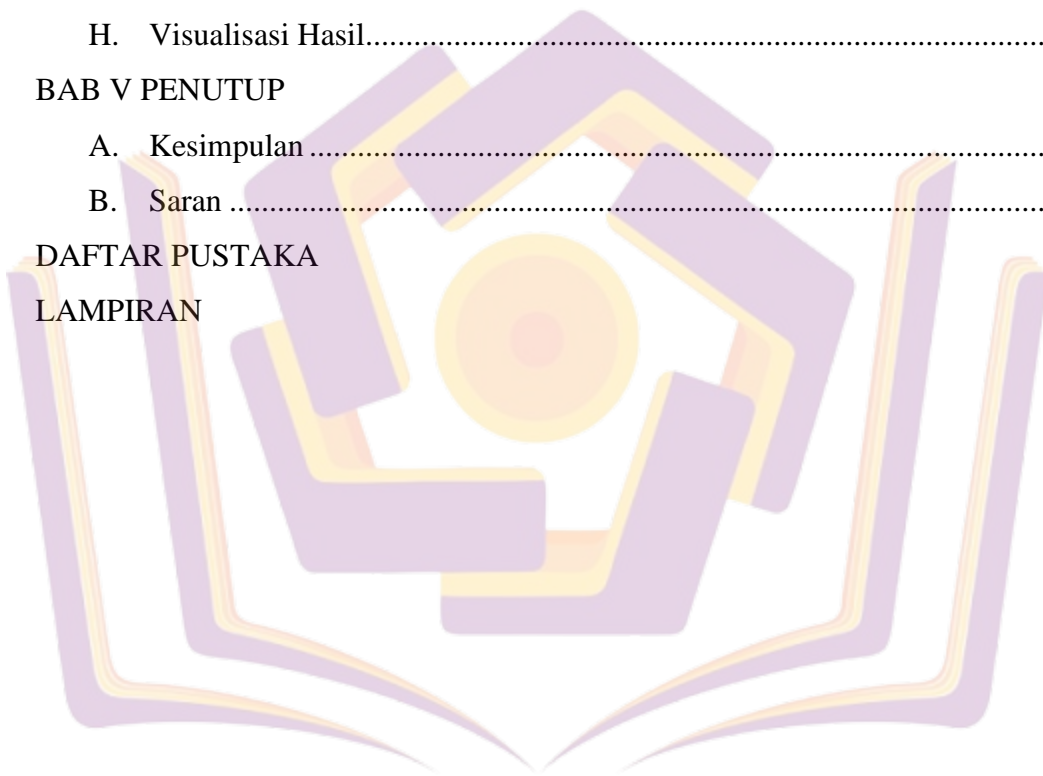
A. Pengumpulan Data .....	28
B. <i>Preprocessing</i> .....	29
C. Pelabelan data .....	37
D. Pembobotan Kata <i>TF-IDF</i> .....	42
E. SMOTE.....	44
F. Klasifikasi Support Vector Machine.....	46
G. Confusion Matrix .....	49
H. Visualisasi Hasil.....	59

## BAB V PENUTUP

A. Kesimpulan .....	62
B. Saran .....	63

## DAFTAR PUSTAKA

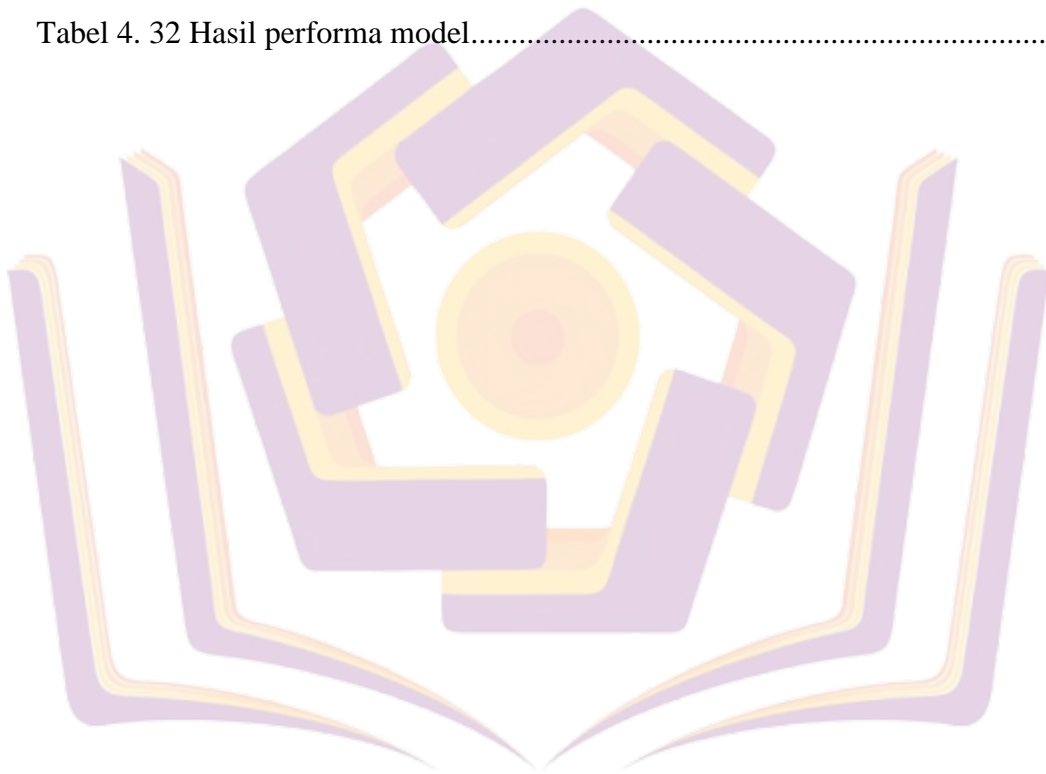
## LAMPIRAN



## DAFTAR TABEL

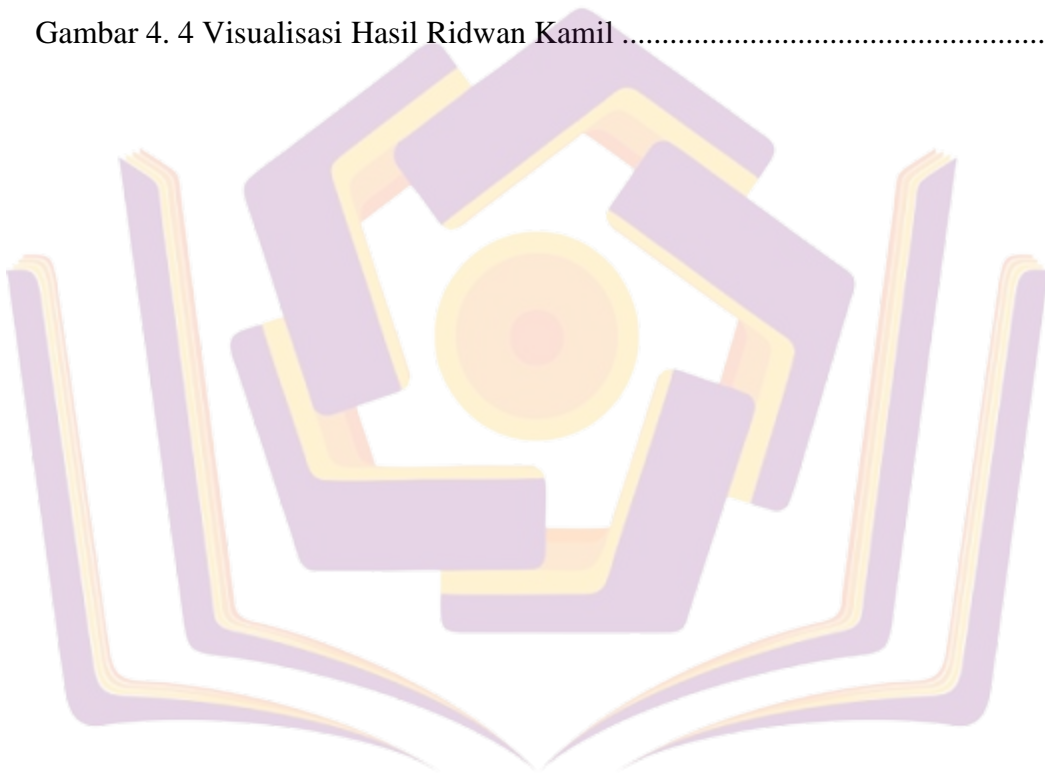
Tabel 2. 1 Confusion Matrix .....	12
Tabel 2. 2 Penelitian Sebelumnya.....	16
Tabel 3. 1 Perangkat Keras ( <i>hardware</i> ) .....	18
Tabel 3. 2 Perangkat Lunak ( <i>software</i> ) .....	19
Tabel 3. 3 Data Hasil <i>Scraping</i> X Anies Baswedan.....	19
Tabel 3. 4 Data Hasil <i>Scraping</i> X Ridwan kamil.....	20
Tabel 4. 1 Data Anies baswedan .....	28
Tabel 4. 2 Data Ridwan Kamil.....	29
Tabel 4. 3 Hasil <i>Cleaning</i> data Anies Baswedan .....	30
Tabel 4. 4 Hasil <i>Cleaning</i> data Ridwan Kamil .....	30
Tabel 4. 5 Hasil <i>Case Folding</i> data Anies Baswedan .....	31
Tabel 4. 6 Hasil <i>Case Folding</i> data Ridwan Kamil .....	31
Tabel 4. 7 Hasil Normalisasi data Anies Baswedan .....	32
Tabel 4. 8 Hasil Normalisasi data Ridwan kamil.....	33
Tabel 4. 9 Hasil <i>Tokenizing</i> data Anies Baswedan .....	33
Tabel 4. 10 Hasil <i>Tokenizing</i> data Ridwan kamil .....	34
Tabel 4. 11 Hasil <i>Filtering</i> data Anies Baswedan.....	35
Tabel 4. 12 Hasil <i>Filtering</i> data Ridwan Kamil .....	35
Tabel 4. 13 Hasil <i>Stemming</i> data Anies Baswedan .....	36
Tabel 4. 14 Hasil <i>Stemming</i> data Ridwan Kamil .....	37
Tabel 4. 15 penghitungan pelabelan lexicon based.....	38
Tabel 4. 16 Hasil <i>labeling</i> data Anies Baswedan.....	38
Tabel 4. 17 Hasil <i>labeling</i> data Ridwan kamil.....	39
Tabel 4. 18 Hasil <i>Labeling</i> Positif.....	39
Tabel 4. 19 Hasil <i>Labeling</i> Negatif .....	40
Tabel 4. 20 Hasil <i>Labeling</i> data Netral .....	40
Tabel 4. 21 <i>Labeling</i> data Anies Baswedan .....	41
Tabel 4. 22 <i>Labeling</i> data Ridwan Kamil .....	41
Tabel 4. 23 Contoh data <i>TF-IDF</i> .....	42

Tabel 4. 24 Proses perhitungan <i>TF-IDF</i> .....	43
Tabel 4. 25 Proses perhitungan bobot ( <i>Weight</i> ) .....	43
Tabel 4. 26 perbandingan data <i>SMOTE</i> Anies Baswedan.....	45
Tabel 4. 27 perbandingan data <i>SMOTE</i> Ridwan Kamil.....	45
Tabel 4. 28 Perbandingan data setelah <i>SMOTE</i> .....	45
Tabel 4. 29 Pembagian data <i>training</i> dan data <i>testing</i> .....	46
Tabel 4. 30 Hasil Confusion Matrix Anies Baswedan.....	50
Tabel 4. 31 Hasil Confusion Matrix Ridwan Kamil .....	55
Tabel 4. 32 Hasil performa model.....	59



## DAFTAR GAMBAR

Gambar 2. 1 Pemisahan kelas SVM.....	8
Gambar 2. 2 Logo Python.....	10
Gambar 3. 1 Konsep Penelitian.....	20
Gambar 4. 1 Confusion Matrix Anies Baswedan .....	50
Gambar 4. 2 Confusion Matrix Ridwan Kamil.....	54
Gambar 4. 3 Visualisasi Hasil Anies Baswedan .....	60
Gambar 4. 4 Visualisasi Hasil Ridwan Kamil .....	60



## DAFTAR LAMPIRAN

Lampiran 1. Sampel dataset Anies Baswedan

Lampiran 2. Sampel dataset Ridwan Kamil

Lampiran 3. Kartu Bimbingan



## DAFTAR ISTILAH

SVM (Support Vector Machine)

NBC (Naïve Bayes Classifier)

TF (*Term Frequency*)

IDF (*Inverse Document Frequency*)

RBF (*Radial Basis Function*)

SMOTE (*Synthetic Minority Over-sampling Technique*)

RAM (*Random Acces Memory*)

