

**PENERAPAN TEXT MINING UNTUK PENCARIAN ARSIP
PADA KELURAHAN KEMANGGISAN
(Studi Kasus : Kelurahan Kemanggisan)**

SKRIPSI

Program Studi Teknik Informatika



Oleh :

Nama : Cipto Febrianto Simanjuntak

Nim : 011 401 503 125 021

**FAKULTAS TEKNIK
UNIVERSITAS SATYA NEGARA INDONESIA
JAKARTA
2019**

**APPLICATION OF TEXT MINING FOR SEARCHING
ARCHIVES IN KEMANGGISAN VILLAGE**

(Study Case : Kelurahan Kemanggisan)

SKRIPSI

Program Of Study ENGINEERING INFORMATICS



Oleh :

Nama : Cipto Febrianto Simanjuntak

Nim : 011 401 503 125 021

**THE FACULTY OF ENGINEERING
UNIVERSITAS SATYA NEGARA INDONESIA
JAKARTA**

2019

**PENERAPAN TEXT MINING UNTUK PENCARIAN ARSIP
PADA KELURAHAN KEMANGGISAN
(Studi Kasus : Kelurahan Kemanggisan)**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar
SARJANA TEKNIK
Program Studi Teknik Informatika**



Oleh :

Nama : Cipto Febrianto Simanjuntak

Nim : 011 401 503 125 021

**FAKULTAS TEKNIK
UNIVERSITAS SATYA NEGARA INDONESIA
JAKARTA**

2019

ABSTRAK

Kelurahan kemanggisan merupakan salah satu kelurahan yang berada di Jakarta barat yang melakukan kegiatan administrasi pemerintahan yang salah satu kegiatannya adalah pengarsipan surat yang ada di kelurahan. Penelitian ini difokuskan untuk melakukan pencarian arsip surat yang ada dalam kelurahan kemanggisan yang melakukan pencarian dengan manual yang memerlukan waktu untuk pencarian arsip surat. *Text mining* merupakan metode klasifikasi variasi data *mining* yang berusaha menemukan pola yang menarik dari data sekumpulan data tekstual yang berjumlah besar. Sedangkan algoritma *naïve bayes* merupakan algoritma pendukung untuk melakukan klasifikasi. Dalam penelitian ini data yang digunakan terdiri dari empat kategori yaitu pemerintahan, ekonomi, keamanan, dan kesejahteraan. Dari setiap kategori terdiri dari beberapa kata kunci (*keyword*). System ini menggunakan Bahasa pemrograman *PHP* dan *database MYSQL* yang menunjukkan bahwa testing bias terklasifikasi secara otomatis. Dari hasil perhitungan probabilitas diketahui bahwa probabilitas surat 11 terhadap kelas ekonomi memiliki nilai paling tinggi, yaitu $3.24223E-60$. Sehingga surat 11 masuk kedalam kelas ekonomi.

Kata kunci : pencarian, klasifikasi, text mining, naïve bayes

ABSTRACT

The village is one of the subdistricts of kemanggisan residing in West Jakarta which performs administrative activities that are one of its activities is the archiving of mail that there is dikelurahan ektp. This study is focused to search archive epistle that exists in the kemanggisan village doing a search with the manual that require time to search the archives. Text mining is a method of classification of variations in data mining are trying to find a pattern that pulls data from a group of large textual data. While the naïve bayes algorithm is algoritman supporters to perform classification. In this study data used consisted of four categories, namely Government, economy, security, and well-being. From each category is composed of multiple keywords (keyword). This system using the programming language PHP and the MYSQL database which showed that testing bias classified automatically. From the results of the calculation of probabilities in mind that probablitas letter 11 against economy class have the highest value, i.e. $3.24223 \text{ E-}60$. So the letter 11 in economy class.

Keywords: search, classification, text minig, naïve bayes

