

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Internet sudah menjadi bagian dari hidup pada saat ini, tidak mudah untuk dapat lepas dari internet. “Pada 2017, “eMarketer” memperkirakan pengguna internet di Indonesia bakal mencapai 112 juta orang, mengalahkan Jepang di peringkat ke-5 yang pertumbuhan jumlah pengguna internetnya lebih lambat”[1]. Dengan perkembangan yang begitu pesat, internet berubah menjadi sumber informasi yang dengan mudah dapat diketahui.

Hoax identik dengan kebohongan atau kepalsuan. Pengertian *Hoax* menurut *Cambridge English Dictionary* adalah “rencana untuk menipu seseorang” misalnya mengatakan kepada polisi ada bom di suatu tempat padahal tidak ada; menurut “*Merriam-Webster*” adalah trik/siasat agar orang percaya atau menerima sesuatu sebagai yang asli padahal palsu dan sering tidak masuk akal[2].

Perwujudan *Hoax* tidak hanya dalam bentuk tulisan, bahkan muncul dalam foto, suara atau bentuk video. Beberapa pengguna internet di Indonesia ada yang kritis dan tidak sedikit yang percaya, berita yang di sampaikan sebagai kebenaran.

Menurut survei yang dilakukan oleh “MASTEL” bahwa 90,30% responden menjawab berita *Hoax* adalah berita bohong yang di sengaja, 61,60% berpendapat berita yang menghasut, 59% berpendapat berita yang akurat, 14% berpendapat berita ramalan dan fiksi ilmiah, 12% berpendapat berita yang

menyudutkan pemerintahan, 3% menjawab berita yang tidak disukai, dan 0.60% berpendapat tidak tahu.[3]

Menteri komunikasi dan informatika mengatakan “sejauh ini sudah ada hampir 800 ribu situs yang menyebar *Hoax* di internet.”[4]. Dari banyaknya jumlah situs yang menyebar *Hoax* di internet maka akan sangat sulit untuk dapat mengelompokkan informasi *Hoax* dan tidak di internet.

Dengan adanya teknologi *Text Mining* yang dapat membantu memecahkan masalah kompleks, didukung perangkat-perangkat teknologi yang semakin maju dan sudah lumrah di gunakan oleh masyarakat umum. Hal tersebut dapat dimanfaatkan untuk memecahkan masalah yang di hadapi pengguna internet terhadap informasi palsu(*Hoax*).

Maka diangkatlah sebuah penelitian skripsi yang berjudul “Klasifikasi Artikel *Hoax* menggunakan *Support Vector Machine Linear* dengan Pembobotan *Term Frequency-Inverse Document Frequency*”.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah di atas, maka masalah yang muncul yaitu: Bagaimana cara merancang sistem yang dapat melakukan klasifikasi informasi dengan melewati bidang keilmuan *Text Mining*.

1.3 Batasan Variabel Masalah

Dalam pelaksanaan penelitian ini, berbagai permasalahan yang muncul dalam konteks objek yang lebih luas akan di batasi sesuai dengan kebutuhan dan kemampuan. Dalam hal ini batasan masalah dalam penelitian ini antara lain:

1. Klasifikasi menggunakan metode *SVM Linear*.
2. Pembobotan *text* menggunakan metode *TF-IDF*.
3. Penelitian ini menggunakan Weka sebagai *tools* untuk mencari model terbaik.

1.4 Maksud dan Tujuan Penelitian

Adapun maksud dan tujuan yang ingin di capai dari penelitian ini adalah menemukan model yang cocok untuk melakukan klasifikasi informasi *Hoax* atau tidak *Hoax* dan sekaligus membuat sistem yang akan menggunakan model tersebut untuk melakukan klasifikasi pada informasi.

1.5 Metode Penelitian

Skripsi ini terdapat metode penelitian yang digunakan untuk mendapatkan informasi yang benar-benar dimengerti dan hasilnya sesuai dengan hasil yang diharapkan serta mendapatkan hasil skripsi yang berkualitas dalam penyusunan laporan, maka menggunakan tahapan-tahapan berikut:

1.5.1 Metode Pengumpulan Data

Metode yang digunakan untuk mengumpulkan data-data dan informasi yang diperlukan untuk membangun sebuah sistem perangkat lunak dalam skripsi adalah:

1. Studi Literatur

Tahapan ini dilakukan dengan mempelajari maupun mengambil kajian dari buku, jurnal, maupun artikel internet untuk dianalisis dan mampu dijadikan bahan pendukung skripsi penelitian yang berkaitan dengan permasalahan yang akan diteliti serta sekaligus sebagai tambahan referensi.

2. Metode Observasi

Merupakan metode dengan melakukan pengamatan langsung di internet untuk mengumpulkan data-data informasi *Hoax* yang beredar di dunia maya. Aktivitas pengumpulan data terkadang terhambat karena sulitnya membedakan informasi *Hoax* atau tidak. Maka penelitian ini mendapatkan informasi *Hoax* dari forum situs yang membahas tentang informasi *Hoax* untuk membantu mendapatkan data.

1.5.2 Metode Analisis Data

Beberapa tahun terakhir mulai populer penelitian data menggunakan metode klasifikasi data. Klasifikasi data erat hubungannya dengan "Data Mining". Sedangkan "Data Mining" adalah proses menganalisis data dari perspektif yang berbeda dan menyimpulkannya menjadi informasi-informasi penting. Secara teknis "Data Mining" dapat diartikan "*data mining refers to extracting or "mining"*

knowledge from large amounts of data. The term is actually a misnomer. Remember that the mining of gold from rocks or sand is referred to as gold mining rather than rock or sand mining, data mining should have been more appropriately named knowledge mining from data”[5]

Data berupa *corpus-corpus* kata yang telah diberi bobot terhadap informasi *Hoax* dan yang tidak, pembobotan dilakukan dengan menggunakan metode TF-IDF. *Corpus-corpus* kata ini didapat dari kumpulan informasi *Hoax* dan informasi yang tidak *Hoax*. Kemudian informasi diberikan bobot terhadap *Hoax* dan tidak *Hoax* dengan melakukan perhitungan jumlah bobot *corpus Hoax* dan tidak *Hoax* sehingga membentuk suatu *dataset* yang akan dijadikan data *training* yang akan di hadapkan dengan metode *SVM* Linear.

1.5.3 Metode Perancangan Perangkat Lunak

Perancangan (*designing*) adalah tahapan penerjemah dari keperluan data yang telah dikumpulkan atau penggambaran perangkat lunak yang akan dibangun. Proses-proses yang dilakukan dalam tahapan perancangan ini terdiri dari:

1.5.3.1 Perancangan Sistem

Metode perancangan sistem dalam penelitian ini menggunakan UML (*Unified Modeling Language*) yaitu himpunan struktur dan Teknik untuk pemodelan desain program berorientasi objek (OO) serta aplikasi meliputi *Use Case Diagram*, *Activity Diagram*, *Sequence Diagram*, *Class Diagram*.

1.5.3.2. Perancangan Database

Metode perancangan *database* dalam penelitian ini adalah metode ERD (*Entity Relationship Diagram*) yaitu gambar atau diagram yang menunjukkan informasi dibuat, disimpan dan digunakan dalam sistem bisnis.

1.5.4 Pengembangan Perangkat Lunak

Tahap ini dilakukan proses mengidentifikasi dan menerjemahkan dari keperluan data atau pemecahan masalah yang telah dirancang ke dalam Bahasa pemrograman komputer. Aktivitas lain yang dilakukan dalam tahapan ini adalah pengujian (*testing*) dan penerapan perangkat lunak dalam lingkungan sistem yang telah diidentifikasi sebelumnya.

1.5.5 Implementasi dan Pengujian Sistem

Tahapan ini aplikasi yang sudah selesai dibuat akan diintegrasikan dan kembali diuji sebagai sistem yang lengkap untuk menjamin seluruh persyaratan sistem yang telah terpenuhi. Apabila aplikasi ini mampu dan dilakukan proses uji coba dan berhasil, maka aplikasi ini akan diimplementasikan dan diterapkan, dan jika belum akan dilakukan perbaikan sistem.

Perangkat lunak ini akan menggunakan metode *k-fold cross validation* sebagai metode pengujian keakuratan algoritma *SVM* yang di hasilkan dan pengujian *black box test* dan *white box test* untuk pengujian sistem.

1.6 Sistematika Penulisan

Penyusunan skripsi ini perlu adanya sistematika penulisan yang terdiri dari bagian-bagian yang saling berhubungan sehingga dapat digunakan sebagai acuan pokok untuk penyusunan laporan skripsi antara lain:

BAB I PENDAHULUAN

Bab ini menguraikan tentang latar belakang masalah, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, manfaat penelitian, metode penelitian, dan sistematika penulisan yang digunakan dalam penyusunan skripsi.

BAB II LANDASAN TEORI

Bab ini berisi tentang teori-teori yang mendasari pembahasan secara rinci, dapat berupa definisi atau model matematis yang langsung berkaitan dengan ilmu atau masalah yang diteliti. Landasan teori dalam bab ini dikelompokkan menjadi empat. Yaitu teori dasar mengenai *data mining*, teori dasar mengenai sistem, teori analisis dan perancangan serta tujuan umum yang meliputi tinjauan pustaka dan tinjauan perangkat lunak.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini akan membahas tentang analisis sistem dan perancangan sistem. Analisis sistem dimulai dari melakukan studi pendahuluan, identifikasi masalah, memahami kerja sistem yang ada, dan hasil

analisis. Perancangan sistem meliputi perancangan struktur menu, perancangan basis data, perancangan proses, serta perancangan *interface*.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Bab ini membahas tentang implementasi dari aplikasi yang akan dirancang yang meliputi cara instalasi aplikasi dan pengoperasian aplikasi. Bab ini juga akan dibahas mengenai hasil dari uji coba aplikasi

BAB V PENUTUP

Bab ini berisi kesimpulan yang didapat dari penelitian pembuatan Aplikasi *DHoax's* berbasis android dan tingkat akurasi dari *SVM* terhadap melakukan klasifikasi terhadap artikel *Hoax* dan kekurangan dari penelitian ini.