

BAB I

PENDAHULUAN

1.1 Latar Belakang

Berita memberikan informasi tentang dunia tempat masyarakat tinggal dengan menempatkan informasi penting, aktual, dan berdampak luas di depan pikiran pembaca, sehingga menambah kualitas hidup masyarakat. Berita juga membentuk cara manusia melihat dunia dan diri sendiri. Menurut Anderson (1983) dalam Wahl Jorgensen dan Thomas (2009: 3) yang menyebabkan berita dapat menjadi perekat sosial yang luar biasa penting untuk di konsumsi terhadap informasi tentang peristiwa terkini, baik besar maupun kecil, mengikat manusia dalam sebuah (*imagined community*).

Terkait permasalahan yang terjadi di Indonesia, baik dalam ranah sosial, budaya, politik, maupun ekonomi, akan memberikan suatu informasi yang nantinya menjadi peluang besar bagi jurnalis atau wartawan di Indonesia, jurnalistik tersebut akan mengumpulkan berita dari berbagai permasalahan yang terjadi, sehingga berita tersebut nantinya dapat termuat dalam *platform* resmi sebagai lapak atau kumpulan dari berbagai berita yang ada. Seperti halnya Jatimnet.com, salah satu *platform* di Jawa Timur yang memuat berbagai macam berita teraktual dan *ter-update*. Jatimnet.com adalah *platform* media *online* lokal yang berbasis di Surabaya. Bernaung di bawah jaringan kerja bersama Jatim Network (Jinet) yang beranggotakan kaum muda dari berbagai bidang; jurnalis, kartunis, pegiat literasi, wirausahawan, dan praktisi teknologi informasi. Jatimnet.com menyajikan informasi mendalam untuk meneropong persoalan menjadi lebih luas. Informasi penting tidak disajikan dengan cara sulit. Sebaliknya ada banyak cara mudah dan menyenangkan untuk menyerap informasi.

Jatimnet.com merupakan *platform* media lokal yang berbasis di Surabaya. Meskipun Jatimnet.com berasal dari media lokal, tetapi para jurnalis maupun editor berpikir secara global. Oleh karena itu Jatimnet.com mengusung tiga gagasan utama dalam berkarya yaitu *diversity*, *creativity*, dan kolaborasi. Sementara keberagaman dan kreativitas menjadi pemicu perkembangan berita tersebut, kolaborasi adalah

strategi kunci memenangkan percaturan media di tengah masifnya perkembangan teknologi. (Redaksi Jatimnet.com, 2019). Jatimnet.com memiliki 12 jurnalis di berbagai daerah Jawa Timur yaitu Surabaya, Gresik, Mojokerto, Ponorogo, Jember, Situbondo, Madiun dan Banyuwangi. Setiap jurnalis yang bertugas akan mengirimkan berita melalui email *jatimnet.com*. Sehingga tugas editor memantau setiap ada berita masuk melalui satu *email*, dan akan mengelompokkan berita tersebut sesuai dengan masing-masing kategori beritanya, dikarenakan ada puluhan berita yang diterbitkan setiap harinya dan hanya memiliki 3 orang yang bertugas sebagai editor, sehingga hal tersebut memakan banyak waktu untuk pengelompokan kategori tiap berita yang masuk. Oleh karena itu dalam penelitian ini akan mengklasifikasikan kategori-kategori berita yang nantinya akan menjadi pengembangan tingkat lanjut ke dalam sistem oleh tim teknologi informasi *jatimnet.com*.

Pada penelitian sebelumnya oleh Shahi, dkk., (2018) melakukan klasifikasi berita Nepal yang memiliki 20 kategori berbeda dan total 4964 dokumen berita dengan menggunakan 3 metode klasifikasi yaitu *Naive Bayes*, *SVM*, dan *Neural Network*. Hasil klasifikasi menunjukkan bahwa *SVM* dengan *RBF* kernel dan *Linear* kernel memiliki akurasi yang lebih tinggi dari 2 metode lainnya yaitu sebesar 74,65% untuk *RBF* kernel diikuti dengan akurasi 74,62% untuk *linear* kernel. Sedangkan untuk 2 metode lainnya memiliki akurasi sebesar 72,99% untuk *Neural Network* dan 68,31% untuk *Naive Bayes*. Selanjutnya pada penelitian Ariadi dan Fithriasari, (2015) tentang Klasifikasi Berita Indonesia menggunakan Metode *NBC* dan *SVM* menghasilkan ketepatan klasifikasi sebesar 88,1%, hasil penelitian menunjukkan bahwa *SVM* kernel *Linear* dan kernel *RBF* menghasilkan ketepatan klasifikasi yang sama dan bila dibandingkan dengan *NBC*, maka *SVM* lebih baik. Menurut Chauhan, dkk., (2018) tentang perumusan dan pemecahan masalah dalam *SVM Linear*. Dalam penelitian ini, *SVM* adalah teknik klasifikasi berbasis margin yang optimal dalam *machine learning*. *SVM* adalah klasifikasi linear yang telah diperluas ke *non-linear* data dengan menggunakan kernel dan data *multi-class* untuk klasifikasi yang lebih dari 2 kelas dengan menggunakan teknik *One vs One* atau *One vs Rest*. *SVM* dengan *Linear* kernel merupakan teknik yang efisien untuk

dimensi tinggi seperti klasifikasi dokumen. Menurut Djajadinata, dkk., (2020) tentang Evaluasi *Feature Extraction* pada Klasifikasi Berita Indonesia, hasil penelitian menunjukkan dengan menggunakan *Gaussian Naive Bayes* menghasilkan akurasi tertinggi sebanyak 99,7% dibandingkan dengan K-NN, *Decision Tree*, *Neural Network*, *Logistic Regression*, dan SVM. Masing-masing memiliki akurasi sebanyak 97,1%, 92,6%, 99,1%, 98,7%, 99%.

Oleh karena itu pada penelitian ini bertujuan untuk mengklasifikasi kategori berita di *website* jatimnet.com menggunakan metode SVM dengan *linear* kernel dan teknik *multi-class* (*One vs One/One vs Rest*) dibandingkan dengan metode *Gaussian Naive Bayes*.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, rumusan masalah dalam penelitian ini adalah bagaimana mengimplementasikan metode *Support Vector Machine* (SVM) dalam klasifikasi berita pada *website* jatimnet.com.

1.3 Tujuan

Berdasarkan rumusan masalah diatas, maka tujuan yang ingin dicapai pada penelitian ini adalah bagaimana mengimplementasi SVM dalam klasifikasi berita yang ada di *website* jatimnet.com.

1.4 Manfaat

Adapun manfaat yang diharapkan dari penelitian ini yaitu:

1. Mempermudah klasifikasi berita pada jatimnet.com sehingga dapat membantu staf/karyawan membangun sistem terbaru dalam menentukan kategori berita kedepannya.

1.5 Batasan Masalah

Adapun batasan masalah pada penelitian ini sebagai berikut:

1. Menggunakan 15 kategori berita (Daerah, Berita, Hukum, Politik, Nasional, Ekbis, Olahraga, Pendidikan, Kesehatan, Gaya Hidup, Tekno, Pilkada,

Inforial, Dunia, dan Ekonomi).

2. Metode atau model yang digunakan untuk proses klasifikasi kategori berita pada *website* jatimnet.com adalah *Support Vector Machine* (SVM).
3. Bahasa yang digunakan pada dataset adalah bahasa indonesia.
4. Ruang lingkup berita berasal dari sumber jatimnet.com.
5. Menggunakan bahasa pemrograman Python dengan *library Scikit-Learn*.