

BAB I

Pendahuluan

1.1 Latar Belakang

Saat ini, emosi manusia merupakan hal yang sering menjadi perhatian dari banyak peneliti pada berbagai macam bidang. Hal ini dikarenakan emosi memiliki peran penting dalam kehidupan sehari - hari manusia, yang menjadikan kebutuhan akan deteksi emosi semakin meningkat. Deteksi emosi manusia dapat berpotensi untuk menyebabkan perkembangan pada berbagai macam area kehidupan seperti hiburan, pembelajaran dan banyak lainnya (Ansari-Asl, et al., 2007) (Liu, et al., 2010).

Pada kehidupan sosial sehari - hari, manusia mengekspresikan berbagai macam emosi yang kompleks melalui berbagai macam cara seperti suara, wajah, tubuh dan fisiologis (sebagai contoh keringat). Manusia dapat dengan mudah mengetahui informasi afektif seseorang dengan menggunakan berbagai macam cara tersebut. Namun, ketika mesin mencoba untuk mengetahui informasi afektif seseorang, dibutuhkan kombinasi dari proses canggih dari metode pemrosesan sinyal (ekstraksi informasi yang relevan dari suatu rekam data) dan algoritma *machine learning* (Ringeval, et al., 2015).

Untungnya, pada saat ini terdapat cukup banyak alat yang dapat digunakan untuk mengekstraksi data fisiologis seseorang seperti *Galvanic Skin Response* (GSR), *Electrocardiogram* (ECG), *electroencephalograph* (EEG), dan lain - lain. GSR merupakan alat yang dapat menghasilkan sinyal fisiologis yang merupakan refleksi dari ketertarikan manusia terhadap suatu hal, ketika manusia tertarik maka akan mengeluarkan keringat sehingga meningkatkan kadar garam dan menambah hambatan

listrik di kulit (AYATA, et al., 2017). Sehingga, GSR bisa dikorelasikan dengan stres dan banyak penelitian yang mendukung hubungan antara ini. Percobaan juga menunjukkan bahwa GSR bisa digunakan untuk menyimpulkan kesulitan atau beban kognitif pada suatu tugas tertentu. Contoh dari kegunaan data GSR ini adalah untuk mengukur performa dari seorang pekerja dalam tingkat stres yang tinggi (Mundell, et al., 2016).

Proses komputasi yang berhubungan dengan, timbul dari atau dipengaruhi emosi disebut dengan *Affective Computing*, yang pertama kali diperkenalkan oleh Picard pada tahun 1997. Tujuan dari *Affective Computing* ini adalah untuk memberikan "emosi" pada komputer sehingga bisa memberikan suatu keputusan, karena komputer yang tidak memiliki "emosi" cenderung untuk berperilaku tidak kreatif dan cerdas. Salah satu cara untuk melakukan *Affective Computing* adalah dengan melakukan proses klasifikasi, karena hal tersebut bisa menjadi permasalahan dalam *Supervised Learning* (Picard, 1997).

Klasifikasi merupakan bagian dari *Supervised Learning*, dimana menggunakan suatu *input* untuk melakukan prediksi nilai dari suatu *output* kualitatif. Proses klasifikasi dapat dilakukan dengan memberikan label kelas pada *output* yang disebut sebagai *target* (Hastie, et al., 2001). Terdapat berbagai macam algoritma untuk melakukan klasifikasi seperti *Support Vector Machine* (SVM), *Naive Bayes*, *K-Nearest Neighbors* (KNN), *Decision Tree* dan banyak lainnya. Penggunaan algoritma tersebut sebaiknya disesuaikan dengan bentuk data yang ingin diklasifikasikan agar mendapatkan hasil yang maksimal.

Akan tetapi seringkali data belum siap untuk diklasifikasi sehingga membutuhkan proses tambahan agar data menjadi siap, data yang sudah siap yaitu data yang

memiliki *input* dan *target output* yang jelas (Famili, et al., 1997). Dikarenakan data GSR merupakan data yang bersifat *univariate* atau hanya memiliki satu variabel, maka diperlukan proses tambahan agar data tersebut siap untuk dilakukan proses klasifikasi. Dalam penelitian ini, data GSR diproses dengan menggunakan metode *sliding windows* agar siap untuk dilakukan proses klasifikasi.

Metode *sliding windows* mengubah permasalahan *sequential supervised learning* menjadi permasalahan *classical supervised learning*. Suatu data sekuensial, sebagai contoh data *univariate* dipetakan ke dalam *windows* sehingga bisa diproses menggunakan algoritma *supervised learning*. Keunggulan yang paling nampak dengan menggunakan metode *sliding windows* adalah dapat diterapkannya semua algoritma *classical supervised learning* serta teknik ini mampu memberikan performa yang memadai pada berbagai macam pengaplikasiannya (Dietterich, 2002).

Dalam penelitian sebelumnya, banyak peneliti menggunakan algoritma SVM dalam mengklasifikasikan data yang diolah dengan metode *sliding windows*. Hal tersebut menjadi pertimbangan penulis dalam memilih algoritma untuk melakukan klasifikasi pada penelitian ini. SVM merupakan algoritma yang dikembangkan oleh Cortes dan Vapnik pada tahun 1995 untuk keperluan klasifikasi biner (Cortes & Vapnik, 1995).

Sampai sekarang SVM merupakan salah satu algoritma yang populer, karena kemampuannya yang hebat dalam klasifikasi *nonlinear*, regresi dan deteksi outlier bahkan sebelum algoritma tersebut populer. Penggunaan SVM lebih mudah jika dibandingkan dengan *Neural Networks*, namun bagi pengguna yang tidak biasa menggunakan SVM akan merasa kesulitan dalam

mengoptimalkan hasilnya. Karena itu, diperlukan suatu pendekatan agar bisa mendapatkan hasil yang baik seperti transformasi, normalisasi, memilih *kernel* yang tepat dan menemukan parameter C dan γ yang tepat (Hsu, et al., 2016).

Secara garis besar SVM bekerja dengan menemukan *hyperplane* yang memiliki margin paling besar pada ruang dimensi yang lebih besar. Penjelasan yang lebih mudah adalah SVM mencoba untuk memisahkan setiap kelas yang ada dengan mencari margin maksimal pada titik kelas terdekat, ketika data tidak bisa diselesaikan dengan pembatas linear maka data akan diproyeksikan pada dimensi yang lebih besar dimana data bisa secara efektif dipisahkan secara *linear* (proyeksi ini dilakukan pada teknik *kernel*). Ada 4 *kernel* dasar pada SVM yaitu linear, *radial basis function* (RBF), *sigmoid* dan *polynomial*. Diantara keempat *kernel* tersebut, *kernel* RBF patut dipertimbangkan untuk dicoba karena kemampuannya dalam memisahkan data *nonlinear* (Meyer, 2017).

Selain SVM, *naive bayes* dan KNN juga cukup sering digunakan sebagai pembanding dalam melakukan klasifikasi pada data yang diolah menggunakan metode *sliding windows*. *Naive bayes* merupakan suatu algoritma klasifikasi yang tidak menggunakan aturan, pohon, atau representasi eksplisit lainnya. Akan tetapi *naive bayes* merupakan pengembangan dari teori probabilitas. Algoritma ini memungkinkan kita untuk menggabungkan beberapa probabilitas pada suatu formula, sehingga bisa digunakan untuk klasifikasi. Sedangkan KNN merupakan algoritma yang biasa digunakan ketika semua atributnya berupa angka, namun algoritma ini dapat dimodifikasi untuk menerima atribut kategori. Cara kerja algoritma ini adalah dengan memperkirakan hasil klasifikasi

menggunakan contoh yang paling dekat dengan data yang ingin diklasifikasikan (Bramer, 2013).

1.2 Rumusan Masalah

Dari latar belakang permasalahan yang ditemui, dihasilkan rumusan masalah sebagai berikut:

1. Bagaimana cara mempersiapkan data *galvanic skin response* menggunakan metode penambangan data temporal sehingga bisa digunakan dalam klasifikasi status emosi manusia dengan baik.
2. Bagaimana menentukan algoritma yang tepat untuk melakukan klasifikasi pada data yang telah dipersiapkan menggunakan metode penambangan data temporal.

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian yang dilakukan adalah sebagai berikut:

1. Mempersiapkan data *galvanic skin response* menggunakan metode penambangan data temporal sehingga bisa digunakan dalam klasifikasi status emosi manusia dengan baik.
2. Menentukan algoritma yang tepat untuk melakukan klasifikasi pada data yang dipersiapkan dengan menggunakan metode penambangan data temporal.

1.4 Batasan Masalah

Batasan-batasan masalah dalam penelitian ini adalah:

1. Penelitian ini berfokus pada pengoptimasian *sliding windows* dan pemilihan algoritma klasifikasi yang sesuai dalam keperluan klasifikasi status emosi manusia agar bisa mendapatkan akurasi yang tinggi.

2. Eksperimen dilakukan dengan menggunakan *setting* laboratorium.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian adalah:

1. Ditemukan cara mempersiapkan data *galvanic skin response* menggunakan metode penambangan data temporal sehingga bisa digunakan dalam klasifikasi status emosi manusia dengan baik.
2. Ditentukan algoritma yang tepat untuk melakukan klasifikasi pada data yang telah dipersiapkan menggunakan metode penambangan data temporal.

1.6 Sistematika Penulisan

Sistematika penulisan pada penelitian ini adalah:

BAB I - Pendahuluan

Pada bab pertama akan dibahas mengenai latar belakang penelitian, rumusan masalah, batasan - batasan penelitian, tujuan penelitian, manfaat penelitian, dan sistematika penulisan penelitian.

BAB II - Tinjauan Pustaka dan Landasan Teori

Pada bab kedua akan dituliskan tinjauan pustaka mengenai penelitian terdahulu mengenai hal yang berkaitan dengan penelitian yang dilakukan saat ini, diikuti dengan landasan teori yang berisi acuan yang digunakan dalam proses penelitian.

BAB III - Metodologi Penelitian

Pada bab ketiga akan dijelaskan mengenai metodologi penelitian yang berisi tentang tahapan - tahapan yang dilakukan dalam proses penelitian.

BAB IV - Hasil dan Pembahasan

Pada bab keempat akan dibahas mengenai data yang didapatkan dari proses eksperimen, dilanjutkan dengan analisis data dan hasil klasifikasi data yang dilakukan.

BAB V - Diskusi

Pada bab kelima akan mendiskusikan secara mendalam mengenai hasil yang didapatkan dari penelitian, yang akan dibandingkan dengan hasil penelitian lain.

BAB VI - Kesimpulan dan Saran

Pada bab keenam akan dijabarkan kesimpulan yang dari penelitian yang dilakukan, diikuti dengan saran untuk penelitian yang akan dilakukan selanjutnya dengan topik bahasan yang serupa.