

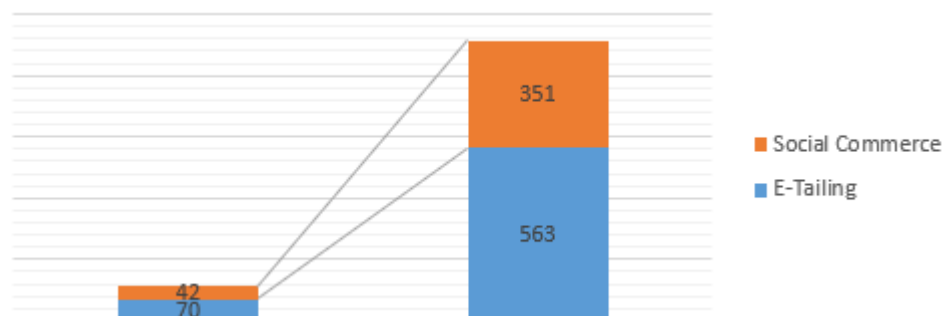
BAB 1

PENDAHULUAN

1.1 Latar Belakang

Masalah yang sering timbul di berbagai negara termasuk di Indonesia salah satunya adalah kemiskinan. Banyak faktor yang menyebabkan kemiskinan itu timbul di negara Indonesia. Pada Bulan Maret 2018, Badan Pusat Statistik mencatat bahwa jumlah penduduk miskin (penduduk dengan pengeluaran per kapita per bulan di bawah Garis Kemiskinan) di Indonesia mencapai 25,95 juta orang (9,82 persen), berkurang sebesar 633,2 ribu orang dibandingkan dengan kondisi September 2017 yang sebesar 26,58 juta orang (10,12 persen)[1].

Di sisi lain, pasar *e-commerce* di Indonesia menjadi pasar *e-commerce* terbesar di Asia Tenggara dengan kontribusi hingga lima puluh persen dari seluruh transaksi di wilayah Asia Tenggara. Peningkatan penetrasi *e-commerce* di Indonesia seiring dengan pertambahan penduduk yang menggunakan internet, sehingga kontribusi terhadap ekonomi Negara Indonesia pada sektor ini berpotensi terus meningkat. Menurut studi McKinsey perusahaan jasa seperti GO-JEK, traveloka, dan platform B2B seperti IndoTrading berada di luar lingkup riset ini. Namun, bahkan tanpa memperhitungkan sektor jasa B2B, *gross merchandise value* pasar *e-commerce* di Indonesia diproyeksikan akan tumbuh sekitar delapan kali lipat pada tahun 2022[2].



Gambar 1-1 Gambar Proyeksi Nilai Pasar *E-commerce* Indonesia (Dalam Triliun Rupiah) [1]

Selama ini data yang didapatkan oleh BPS adalah dengan menyelenggarakan Survey Sosial Ekonomi Nasional yang biasa disingkat Susenas. Susenas adalah survey berbasis rumah tangga yang mengumpulkan informasi mengenai karakteristik sosial ekonomi seperti: pendidikan, kesehatan, keluarga berencana, keterangan bepekerjaan, kriminalitas, perumahan, perlindungan sosial serta konsumsi dan pengeluaran rumah tangga. Susenas pertama dimulai sejak tahun 1963 dan selalu diselenggarakan setiap tahun [3]. Survey yang dilaksanakan oleh BPS yaitu Susenas, dapat diperkirakan akan memakan waktu yang cukup lama dan akan ada rentang waktu antara survey satu dengan yang lainnya, bahkan jika dihitung biaya survey untuk pendataan data masyarakat kurang mampu, dapat dipastikan akan memakan biaya yang cukup besar. Dengan adanya teknologi informasi khususnya teknologi *Machine Learning*, cara survey untuk pendataan kemiskinan di suatu daerah akan lebih mudah dan akurat. Bahkan dengan pengembangan lebih lanjut, teknologi tersebut akan dapat memantau naik turunnya kemiskinan di Indonesia tanpa waktu lama.

Dari permasalahan diatas, dapat ditarik solusi untuk menyelesaikan permasalahan waktu dan biaya survey Susenas yaitu dengan mengembangkan sebuah perangkat lunak dengan teknologi *Machine Learning*. *Machine Learning* yang digunakan akan melengkapi survey kemiskinan pada suatu daerah dengan data dari aplikasi *e-commerce* yang sedang marak digunakan untuk masyarakat Indonesia. Dengan adanya aplikasi tersebut, survey mengenai kemiskinan di suatu daerah di Indonesia akan sangat terbantu dan tidak akan memakan waktu yang lama dalam mendapatkan data karena akan menggunakan data *e-commerce*.

E-commerce yang sering dipakai oleh masyarakat Indonesia dapat dipakai untuk menjadi media prediksi kemiskinan di Indonesia. Menggunakan perhitungan jumlah pembelian dari suatu daerah dapat diketahui di daerah tersebut apakah daerah tersebut termasuk daerah yang Makmur atau sebaliknya. Kemudian dari perhitungan tersebut dapat dijadikan proses prediksi kemiskinan menggunakan aplikasi dengan basis *machine learning*.

1.2 Rumusan Masalah

Rumusan masalah yang mendasari pengembangan Aplikasi *Machine Learning* Menggunakan Algoritma *Support vector machine* dan *Statistical-based Feature Selection* Untuk Memprediksi Kemiskinan adalah sebagai berikut:

1. Bagaimana cara melengkapi hasil survey dan sensus tingkat kemiskinan di suatu daerah di Indonesia berdasarkan data *E-commerce*?
2. Bagaimana cara mengidentifikasi item yang berpengaruh dalam prediksi kemiskinan?
3. Bagaimana cara penyajian data hasil dari prediksi kemiskinan?

1.3 Tujuan

Tujuan yang diharapkan menyelesaikan rumusan masalah pada pembuatan proyek akhir ini adalah dengan sebagai berikut:

1. Untuk melengkapi hasil survey dan sensus kemiskinan adalah dengan menggunakan algoritman *machine learning support vector machine* dan *statistical-based feature selection*.
2. Untuk mengidentifikasi item yang berpengaruh dari data e-commerce adalah dengan menggunakan algoritma seleksi fitur yaitu *statistical-based feature selection*.
3. Untuk menyajikan data hasil prediksi kemiskinan adalah dengan menggunakan platform web yang di dalamnya akan menjalankan algoritma *support vector regression*.

1.4 Batasan Masalah

Batasan masalah dapat berisi:

1. Aplikasi hanya memprediksi dan menyajikan data kemiskinan dari data aplikasi *e-commerce*.

2. *Dataset* yang digunakan adalah *dataset e-commerce* di pulau jawa.
3. Hanya akan terdapat satu pengguna dalam aplikasi ini.
4. Algoritma seleksi fitur yang bisa digunakan pada aplikasi web hanya 3, yaitu: f-score; chi-square; correlation-based feature selection.
5. Aplikasi ini menyediakan fitur konfigurasi 2 *hyperparameter* pada SVR yaitu parameter regularisasi (C) dan nilai epsilon.

1.5 Metode Pengerjaan

Metode pengembangan perangkat lunak yang digunakan pada pengerjaan proyek akhir ini adalah dengan menggunakan SDLC *Waterfall* [2]. Berikut gambaran tahapan – tahapan SDLC *sequential linear*:



Gambar 1-2 Gambar Tahapan SDLC Waterfall [4]

Dari gambaran diatas dapat dijabarkan mengenai metode pengerjaan yang digunakan saat pengerjaan proyek akhir. SDLC *sequential linear* digunakan karena pada pengembangan aplikasi ini menggunakan waktu yang relatif pendek. Lama pengerjaan akan di jelaskan secara detail di bagian jadwal pengerjaan. Berikut adalah penjabaran dari tahapan SDLC *sequential linear*:

1. Analisis

Pada tahap ini dilakukan penganalisaan kebutuhan aplikasi dari mulai penganalisaan *dataset* yang akan digunakan. Selain kebutuhan dari sisi aplikasi, pada tahap ini dianalisis kebutuhan *library* yang akan digunakan terutama pada bagian permodelan *machine learning*.

2. Desain

Pada tahap desain, dilakukan pembuatan desain kebutuhan aplikasi seperti kebutuhan database, alur aplikasi, permodelan alur data *machine learning*, dan desain antarmuka pengguna.

3. Implementasi

Proses implementasi pada proyek akhir ini akan menggunakan platform web sebagai antarmuka pengguna dengan sistem *machine learning* yang digunakan untuk memprediksi kemiskinan. Algoritma *machine learning* yang diimplementasikan pada aplikasi ini adalah algoritma *support vector regression*. Karena *dataset* yang digunakan adalah *dataset* yang berukuran cukup besar, maka proses permodelan *machine learning* akan dibantu dengan metode seleksi fitur dengan algoritma *statistical-based feature selection*.

4. Pengujian

Aplikasi proyek akhir ini akan diuji dengan melakukan pengujian metode *black-box* pada bagian antarmuka pengguna untuk menguji kesesuaian respon aplikasi dengan aktifitas user.

1.6 Jadwal Pengerjaan

Pengembangan aplikasi mulai dari desain dan perancangan hingga ke tahap pengujian diatur pada jadwal pengerjaan. Berikut jadwal pengerjaan yang mengatur waktu pengerjaan aplikasi ini:

Tabel 1-1 Jadwal Pengerjaan Proyek Akhir

Kegiatan	Oktober 2019				November 2019				Desember 2019				Januari 2020				Februari 2020				Maret 2020				April 2020			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Pengumpulan kebutuhan																												
Analisis dan Desain																												

