

BAB 1

PENDAHULUAN

Bab pendahuluan merupakan penjabaran dari latar belakang masalah yang terjadi, rumusan masalah yang dikerucutkan dari masalah, ruang lingkup penelitian, tujuan penelitian, dan sistematika penulisan dari penelitian yang dilakukan.

1.1 LATAR BELAKANG

Manufaktur merupakan istilah bagi kegiatan industri yang memproduksi produk dengan jumlah banyak dari bahan mentah menjadi produk setengah jadi ataupun barang jadi sehingga memiliki nilai jual lebih (ReVelle, 2002). Dalam manufaktur, proses produksi akan menentukan hasil dari kualitas produk yang dihasilkan, dikarenakan semua proses yang berkaitan langsung dengan hasil produk hampir seluruhnya berada di bagian ini (Bechmann et al., 2015). Suatu produk haruslah sesuai dengan yang diharapkan konsumen. Kualitas menjadi sebuah tolak ukur antara industri sejenis sebagai faktor daya saing, pentingnya kualitas ini tentunya akan menambah kepuasan bagi konsumen (Juran et al., 1999). Berjalannya proses produksi sebuah industri manufaktur tidak akan terlepas dari produk dengan kualitas dibawah standar atau dapat disebut produk cacat (Zhang et al., 2020). Produk yang memiliki kualitas dibawah standar perlu diatasi agar tidak menyebabkan kerugian besar baik secara biaya ataupun waktu. Biasanya tindakan yang dilakukan untuk menangani produk cacat tersebut dengan di-*rework* ataupun menjadi produk *reject*.

Untuk mengurangi probabilitas terjadinya produk cacat, Data-data proses produksi dapat dimanfaatkan untuk mengetahui parameter-parameter yang paling berpengaruh kepada produk secara langsung berdasarkan pola yang terbentuk dari data-data yang sudah ada (House & Rado, 2013), sehingga perusahaan dapat memprediksi kualitas produk pada produksi selanjutnya. Prediksi kualitas

merupakan bagian dari manajemen proses perbaikan kualitas yang memudahkan tindakan lanjut dari produk yang mengalami kecacatan (Bechmann et al., 2015). Model prediksi kualitas merupakan alat untuk membantu keputusan perbaikan kualitas tersebut. Pemodelan kualitas yang dilakukan harus objektif dan memiliki hasil prediksi yang lebih akurat daripada tes-tes yang sudah dilakukan (Demirbilek & Gregoire, 2017). Prediksi kualitas dapat dilakukan dengan berbagai teknik, salah satunya menggunakan *data mining*. *Data mining* adalah proses penerapan metodologi komputer dan statistik dengan teknik-teknik baru untuk melakukan penemuan pengetahuan dari data. Standar pendekatan *data mining* yang umum di berbagai sektor industri adalah *cross-industry standard process for data mining* (CRISP-DM). CRISP-DM merupakan kerangka kerja untuk menerjemahkan masalah bisnis ke dalam penambangan data dengan enam fase proses (Huber et al., 2019).

Pengembangan yang dilakukan dari pendekatan CRISP-DM menghasilkan *data mining* yang dikhususkan untuk domain rekayasa, pendekatan ditambahkan unsur teknis yang tidak dijelaskan dalam metode CRISP-DM, hasil dari pengembangan pendekatan tersebut disebut *data mining for engineering application* (DMME). Pendekatan DMME yang menambahkan unsur teknis dalam proses *data mining* dapat dilakukan dengan lebih cepat karena proses dilakukan secara sistematis sehingga dapat dikembangkan dan didokumentasi lebih lanjut (Huber et al., 2019). Prediksi kualitas dilakukan sebelum proses produksi dimulai, agar dapat menekan risiko produk dibawah standar dengan mengandalkan data historis perusahaan.

Penelitian tentang model prediksi kualitas sudah banyak dilakukan. Salah satu penelitian yang pernah dilakukan adalah *wafer dicing quality prediction* membahas mengenai prediksi keabnormalan pada produksi *wafer dicing*. Hasil dari penelitian ini adalah model secara efektif dapat memprediksi kegagalan *wafer* dengan akurasi mencapai 75% (Su et al., 2018). Penelitian dalam bidang lain adalah *coke quality prediction* yang membahas kaitan antara karakteristik *coal* dengan mempertimbangkan parameter *coke strength after reaction* (CSR) dan *coke reactivity index* (CRI) terhadap *coke* yang dihasilkan. Penelitian tersebut

menghasilkan pengetahuan mengenai hubungan dari karakteristik *coal* dengan parameter CSR dan CRI menggunakan pendekatan model regresi (North et al., 2018). Penelitian pada *battery quality prediction* yang membahas tentang rantai proses yang rumit sehingga mengalami berbagai kendala produksi. Penelitian berfokus pada model regresi multivariat dan menghasilkan model prediksi untuk *multi-critical final product properties* (FPP) dengan pendekatan *data-driven*. Penelitian lebih lanjut dapat meliputi metode lain seperti *neural network*, *decision tree*, dan aturan asosiasi. Percobaan selanjutnya dapat difokuskan pada pemahaman efek parameter perubahan pada implementasi ke dalam kontrol mesin secara langsung (Thiede et al., 2019).

Salah satu penelitian lain berkaitan dengan prediksi kualitas pada kasus dataset manufaktur semikonduktor yaitu *cascade quality prediction method using multiple PCA+ID3 for multi-stage manufacturing system* yang membahas cara untuk mencapai manufaktur tanpa cacat dan pengembangan model prediksi kualitas dalam *multi-stage manufacturing* (MMS). Penelitian menghasilkan performa metode *cascade quality prediction method* (CQPM) untuk kasus dataset semikonduktor yang terbukti memiliki performa lebih baik dari model lainnya. Akurasi pada metode sebesar 90,02% dan *G-mean* sebesar 0,4448 (Arif et al., 2013).

Terdapat juga penelitian sebelumnya yang dilakukan mengenai pembuatan *prototype* untuk menghasilkan *tools machine learning* berbasis *website* (Wildan., 2018). Penelitian tersebut menghasilkan *prototype* dengan model berakurasi 95,64% pada set data percetakan dengan 5 parameter. Akan tetapi, pendekatan yang digunakan masih belum spesifik pada domain rekayasa karena masih menggunakan CRISP-DM, kemudian penelitian tersebut juga belum dapat membuktikan bahwa terdapat hubungan parameter terhadap kualitas akhir produk dan keakuratan prediksi akan sama apabila parameter yang digunakan memiliki nilai konten yang bervariasi di masing-masing atributnya dan memiliki atribut yang lebih banyak sehingga dapat dikatakan bahwa model tersebut belum dapat dipastikan valid untuk dataset lain. Penelitian tersebut juga hanya spesifik untuk memprediksi kualitas pada produk cacat *cover* dan presisi pada *offset printing*

dengan 5 parameter. Apabila terdapat penambahan parameter proses maka algoritma untuk *rules* model perlu diuji kembali sehingga diketahui bahwa model tersebut benar-benar valid memiliki akurasi tinggi dan jika model berubah maka *prototype* harus dibuat ulang.

Dataset semikonduktor dikumpulkan dari proses pemantauan berkelanjutan menggunakan sensor dan peralatan metrologi. Proses pengambilan data pada masing-masing sensor dinomori dengan variabel S1 hingga S590 (McCann & Johnston., 2008). Perekapan data yang dilakukan banyak sensor tersebut menghasilkan banyak data yang dapat dimanfaatkan untuk melakukan prediksi kualitas. Walaupun begitu, teknik-teknik yang dilakukan dalam penelitian-penelitian diatas, belum tentu tepat untuk diterapkan pada dataset semikonduktor yang diteliti pada penelitian ini karena perbedaan karakteristik data. Dampak yang dihasilkan dari penelitian ini yaitu penggunaan *data mining* yang semulanya secara umum hanya dilakukan secara CRISP-DM dapat digantikan dengan cara yang memiliki kekhususan pada domain rekayasa sehingga pengaplikasian *data mining* lebih mengakomodasi kebutuhan *engineering* pada suatu perusahaan dengan menggunakan DMME. Dampak lainnya adalah banyaknya produk yang dapat diperbaiki dan ditingkatkan pada suatu proses produksi.

1.2 RUMUSAN MASALAH

Penelitian-penelitian di atas sebagian besar berfokus pada prediksi kualitas, pemodelan pendekatan terbaru, dan penerapannya pada masalah yang spesifik. Walaupun permasalahan-permasalahan yang sudah diuraikan dapat diatasi dengan menggunakan pendekatan *data mining*, akan tetapi penelitian-penelitian tersebut diterapkan secara spesifik pada satu dataset. Suatu algoritma yang terbukti akurat pada sebuah dataset belum tentu memiliki tingkat akurasi yang sama ketika diterapkan pada dataset yang lain. Hal ini disebabkan beberapa permasalahan seperti perbedaan tipe data dan jumlah variabel yang dapat digunakan. Pada kasus manufaktur semikonduktor yang direpresentasikan oleh dataset SECOM atribut merupakan variabel multivariat dengan tipe data numerik

dengan kelas nominal. Belum diketahuinya teknik yang tepat dan pengaruh parameter proses terhadap kualitas akhir produk pada dataset semikonduktor menjadi dasar dari penelitian ini.

1.3 RUANG LINGKUP PENELITIAN

Ruang lingkup dalam penelitian ini adalah dataset SECOM yang merupakan data sekunder dari data manufaktur semikonduktor yang didapat dari penelitian terdahulu (McCann & Johnston., 2008). Selain itu, terdapat pula asumsi bahwa pendekatan sistem manufaktur pada penelitian ini menggunakan pendekatan *single-stage manufacturing*.

1.4 TUJUAN PENELITIAN

Penelitian ini bertujuan untuk menghasilkan model prediksi dengan tingkat akurasi tinggi pada kasus data dengan atribut numerik dan kelas nominal.

1.5 SISTEMATIKA PENULISAN

Sistematika penulisan menjabarkan secara singkat dari laporan tugas akhir yang dilakukan oleh peneliti pada setiap bab yang ada.

BAB I PENDAHULUAN

Bab ini menjelaskan mengenai hal yang menjadi dasar tugas akhir ini dibuat yaitu perancangan model prediksi kualitas. Hal-hal yang dijelaskan pada bab ini meliputi latar belakang masalah, rumusan masalah, ruang lingkup penelitian, tujuan penelitian dan sistematika penulisan. Latar belakang masalah membahas mengenai permasalahan yang diangkat yaitu tentang model prediksi kualitas untuk kasus dataset semikonduktor. Rumusan masalah menjelaskan permasalahan mengenai teknik dan pengaruh parameter proses terhadap kualitas akhir produk. Tujuan penelitian untuk menghasilkan model prediksi dengan tingkat akurasi tinggi pada kasus data dengan atribut numerik dan kelas nominal. Ruang lingkup penelitian menjelaskan lingkup penelitian yang dilakukan bersamaan dengan asumsi sistem manufaktur yang menggunakan pendekatan *single-stage manufacturing*. Sistematika penulisan memaparkan mengenai

pembahasan secara singkat di masing-masing bab penelitian.

BAB II LANDASAN TEORI

Bab ini menjabarkan studi literatur yang mendukung penelitian, hal yang dijelaskan meliputi Kualitas, *Data Mining*, DMME, *Machine Learning*, Rapidminer dan literatur pendukung lainnya.

BAB III METODOLOGI PENELITIAN

Bab metode penelitian ini menjelaskan tahap atau skema untuk melakukan penelitian yang akan dilakukan untuk memecahkan masalah yang terjadi. Skema divisualisasikan dengan media berupa *flowchart* atau diagram alir. Metodologi yang digunakan pada penelitian ini mengadaptasi dari *data mining methodology for engineering application* (DMME) dengan beberapa penjelasan dari alur penelitian yang diawali dari landasan teori, rumusan masalah, tujuan penelitian, *data understanding*, *data preparation*, *modelling*, *evaluation*, sampai penutup yaitu kesimpulan dan saran.

BAB IV FORMULASI MODEL

Bab formulasi model memaparkan proses *data mining* dan analisis pada dataset yang diteliti yaitu dataset semikonduktor dari awal hingga akhir proses *data mining*. Proses meliputi *data understanding*, *data preparation*, *modelling*, dan *evaluation* sampai mendapatkan *rules model*.

BAB V PENUTUP

Bab ini berisi kesimpulan yang dilakukan dari hasil keseluruhan penelitian, menjelaskan mengenai poin-poin yang dapat disimpulkan dari hasil penelitian yang dilakukan dan saran yang diberikan untuk menunjang penelitian lebih lanjut kedepannya.