

## BAB 2

### TINJAUAN PUSTAKA DAN DASAR TEORI

Bab ini berisi tentang penelitian-penelitian sebelumnya, teori-teori yang dapat menjadi referensi dan dasar dalam penelitian ini, serta perbandingan antara penelitian terdahulu dengan penelitian sekarang.

#### 2.1. Tinjauan Pustaka

Penelitian-penelitian mengenai *text mining* yang telah dilakukan sebelumnya menjadi acuan dalam penulisan penelitian sekarang. Subbab ini menjelaskan tentang penelitian sebelumnya yang memiliki keterkaitan dengan penelitian yang dilakukan oleh penulis. Literatur yang dipelajari dalam penelitian ini terkait dengan penanganan komplain agregat, implementasi *text mining* pada Departemen *Quality Assurance*, penggunaan *text mining* terkait dengan pendapat terhadap produk serta penggunaan *One Simple Sign Test* untuk membandingkan pendapat terhadap sistem pengelolaan komplain di perusahaan saat ini dengan sistem baru yang ditawarkan. Berdasarkan pencarian kata kunci terkait dengan penggunaan *text mining* di Emerald, ditemukan sejumlah 278 dokumen yang telah dipublikasikan seperti terlihat pada Gambar 2.1. Setelah melihat pada dokumen-dokumen yang telah dipublikasikan, belum ditemukan penelitian yang menggunakan *e-mail* sebagai data yang akan digali.



Gambar 2.1. Literature Review

##### 2.1.1. Penelitian Sebelumnya

Ashton dkk (2014) mengaplikasikan *Latent Semantic Analysis* untuk penyusunan *control chart* dengan input berupa opini pelanggan. Pada penelitian tersebut, faktor konsep yang terkait dengan opini pelanggan diekstrak dari data besar pada retailer khusus dalam Fortune 500 yang menawarkan layanan online kepada pelanggan di Amerika Serikat. Data yang digunakan terdiri dari 820.000 lebih komentar pelanggan yang memutuskan untuk membatalkan layanan berlangganan mereka dengan penyedia dalam periode 27 bulan. Dalam proses pembatalan, pelanggan diminta untuk mengisi alasan pembatalan pada kueri yang tersedia. Sebelumnya

penyedia layanan menggunakan cara tradisional untuk menganalisis alasan pelanggan membatalkan layanan, seperti membaca dan menafsirkan tulisan pelanggan.

Sikap dan penilaian publik terhadap kebijakan pemerintah Nigeria yang berlaku saat itu diteliti oleh Adeyemo dan Ojo (2014). Adapun pengumpulan data dilakukan dengan cara mengumpulkan komentar di media sosial seperti Facebook dan Twitter. Komentar-komentar tersebut kemudian diekstraksi ke dalam file Microsoft Excel dengan format CSV. Tiap komentar dalam file tersebut kemudian dipisahkan dengan menggunakan *File-Splitter* dan diolah dengan menggunakan *Rapid Miner* dengan algoritma *K-means Clustering*.

*Text mining* digunakan untuk menemukan dan mengidentifikasi kegagalan yang ditemukan di beberapa armada pesawat yang ada di Amerika, seperti diteliti oleh Rodrigues dkk (2012). Input dari penelitian tersebut merupakan laporan kesalahan atau kegagalan yang ditulis oleh operator maupun laporan permasalahan yang dicatat oleh awak pesawat. Data yang diperoleh kemudian diolah dengan menggunakan *text mining* dan *Artificial Neural Network (ANN)*. Tujuan dari penggunaan *text mining* dalam penelitian tersebut adalah mempercepat dan mempermudah identifikasi kegagalan di armada mereka supaya perusahaan mampu bersaing di dunia penerbangan.

Dalam jurnalnya, Tan (2012) membahas aplikasi *text mining* untuk mendiagnosis kesalahan yang terjadi pada suatu sistem manufaktur. Dalam jurnal tersebut dicontohkan beberapa aplikasi pengelompokan atau klasifikasi dan *clustering*. Klasifikasi adalah fungsi yang sangat berguna di banyak bidang manufaktur. Sebagai contoh, dalam industri semikonduktor, cacat diklasifikasikan untuk menemukan skema yang kemudian akan digunakan untuk meningkatkan kualitas hasil produksi.

Manns dkk (2015) dalam jurnalnya membahas tentang aplikasi *text mining* untuk mengolah data *work task* di *assembly line* Daimler. *Text mining* digunakan untuk meningkatkan efisiensi dengan cara menampilkan rencana kerja yang cocok untuk menghasilkan produk baru. Metode *Controlled Natural Language* dikembangkan untuk mendukung perencanaan proses yang dapat mempersingkat waktu pengembangan produksi.

*Text mining* dapat digunakan untuk meningkatkan proses *New Product Development (NPD)* yaitu dengan mengidentifikasi masalah yang sering muncul

pada proses NPD, seperti diteliti oleh Costa dkk (2013). Data yang digunakan adalah data hasil wawancara dengan personel NPD dan data hasil dari metode diagnostik *Current Reality Tree* (CRT) yang digunakan untuk mengidentifikasi efek yang tidak diinginkan dalam suatu proses. Setelah itu, masalah yang diperoleh dikelompokkan untuk selanjutnya dilakukan analisis terhadap permasalahan tersebut.

### 2.1.2. Penelitian Sekarang

Penelitian dilakukan dengan mengumpulkan *e-mail* berisi komplain yang dikirimkan oleh pelanggan terkait dengan kualitas produk di PT. X Indonesia. Seperti diketahui di mana sebelumnya Supervisor *Quality Assurance* menggunakan cara tradisional dengan cara memberi tanda pada bagian dari *e-mail* yang merupakan "kata kunci" dari kesalahan yang menyebabkan pelanggan melayangkan komplain. Peneliti menggunakan tahapan *Text Preprocessing* dan melanjutkannya dengan mencocokkan hasil *Text Preprocessing* dengan data masa lalu untuk menemukan permasalahan yang termuat dalam *e-mail* komplain, kemudian mengklasifikasikan permasalahan yang terdapat dalam seluruh *e-mail* komplain, seperti dalam penelitian Costa dkk (2013) yang mengklasifikasikan permasalahan dalam *New Product Development*. Penerapan *text mining* bertujuan untuk meningkatkan efisiensi dalam pengelolaan komplain. Adapun algoritma *text mining* akan dibangkitkan menggunakan Program Visual Studio C#.

Ashton dkk (2014), Adeyemo dan Ojo (2014) menggunakan komentar dari media *online* sebagai input dari penelitian. Rodrigues dkk (2012) dalam penelitiannya menggunakan catatan kegagalan yang ditulis oleh awak pesawat dan operator sebagai inputnya, sedangkan penelitian yang akan penulis lakukan menggunakan *e-mail* berisi komplain yang diterima dari pelanggan, sedangkan Ashton dkk (2014) dan Rodrigues dkk (2012) dalam penelitiannya menggunakan metode *Latent Semantic Analysis*, Adeyemo dan Ojo (2014) menggunakan metode *K-means Clustering*. Penelitian ini merupakan penelitian untuk penyelesaian suatu masalah atau *Problem Solving Research* di mana tujuan penelitian ini adalah untuk meningkatkan atau memperbaiki sistem yang berlaku saat ini supaya lebih efisien. Seperti pada penelitian Manns dkk (2015) yang ingin meningkatkan efisiensi dalam menyusun rencana proses untuk menghasilkan produk baru. Perbandingan antara penelitian sekarang dengan penelitian terdahulu dapat dilihat pada Tabel 2.1.

**Tabel 2.1. Perbandingan Penelitian Sekarang dengan Penelitian Terdahulu**

<b>Penulis</b>	<b>Judul</b>	<b>Data</b>	<b>Metode</b>	<b>Objek</b>	<b>Tujuan</b>
Ashton dkk (2014)	<i>Quantitative quality control from qualitative data: control charts with latent semantic analysis</i>	Kueri <i>online</i> alasan pelanggan membatalkan / berhenti berlangganan	<i>Latent Semantic Analysis</i>	<i>Retailer</i> yang termasuk dalam Fortune 500	Menyusun <i>control chart</i> berdasarkan opini pelanggan
Adeyemo dan Ojo (2014)	<i>Classification of Social Blogs Comments Using Text mining</i>	Komentar masyarakat di Facebook dan Twitter tentang kebijakan pemerintah Nigeria	<i>K-mean Clustering</i>	Kebijakan Pemerintah Nigeria	Mengetahui respon masyarakat terhadap kebijakan pemerintah Nigeria
Rodrigues dkk (2012)	<i>Aircraft interior failure pattern recognition utilizing text mining and neural networks</i>	Catatan kegagalan yang ditulis oleh awak pesawat dan operator	<i>Latent Semantic Analysis</i> dan <i>Artificial Neural Network</i>	Beberapa Armada pesawat di Amerika	Menyusun langkah perbaikan berdasarkan kegagalan yang ditemukan

**Tabel 2.1. Lanjutan**

<b>Penulis</b>	<b>Judul</b>	<b>Data</b>	<b>Metode</b>	<b>Objek</b>	<b>Tujuan</b>
Tan (2012)	<i>Application of Data Mining in Fault Diagnosis</i>	Berbagai data yang terkait dengan kegagalan	Klasifikasi, <i>clustering</i>	Berbagai sistem manufaktur	Menunjukkan contoh aplikasi <i>text mining</i> untuk diagnosis kesalahan
Manns dkk (2015)	<i>Automatic proposal of assembly work plans with a controlled natural language</i>	Data work task di <i>assembly line</i> Daimler	<i>Controlled Natural Language</i>	<i>Assembly line</i> Daimler	Meningkatkan efisiensi dalam menyusun rencana kerja yang cocok untuk produk baru
Costa dkk (2013)	<i>Systematization of Recurrent New Product Development Management Problems</i>	Data hasil wawancara dengan personel <i>New Product Development</i>	Mencocokkan hasil wawancara dengan data hasil dari metode diagnostik <i>Current Reality Tree (CRT)</i>	Sistem <i>New Product Development</i>	Mendiagnosis dan mengklasifikasikan permasalahan dalam <i>New Product Development</i>

**Tabel 2.1. Lanjutan**

<b>Penulis</b>	<b>Judul</b>	<b>Data</b>	<b>Metode</b>	<b>Objek</b>	<b>Tujuan</b>
Penulis (2018)	Penggunaan <i>Text Mining</i> untuk Pengelolaan <i>E-mail</i> Komplain di PT. X Indonesia	<i>E-mail</i> berisi komplain yang dikirimkan oleh pelanggan ke perusahaan	<i>Text Preprocessing</i> dilanjutkan dengan pengelompokan cacat	Departemen QA di PT. X Indonesia	Menyusun perbaikan pada proses perekapan <i>e-mail</i> komplain



## 2.2. Dasar Teori

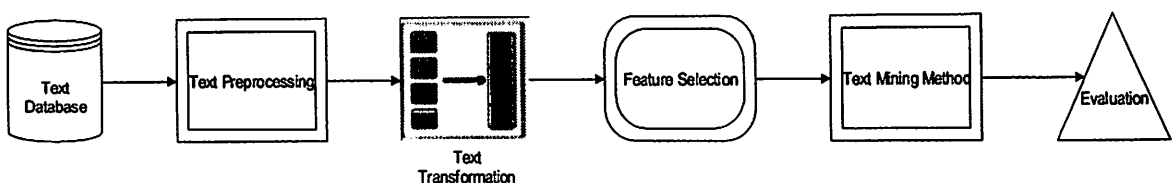
Penelitian Penggunaan *Text Mining* untuk Pengelolaan *E-mail* Komplain di PT. X Indonesia ini ditulis dengan mengacu pada teori-teori maupun temuan-temuan yang ada. Subbab ini berisi teori-teori dan temuan-temuan dari penelitian terdahulu yang menjadi dasar dalam penelitian ini.

### 2.2.1. Penanganan Komplain

Penanganan komplain terdiri dari dua fungsi dasar, yaitu penanganan komplain individual dan analisis dari data komplain agregat (Mitchell, 1993). Untuk mendapatkan manfaat penuh dari komplain, perlu untuk mencatat semua komplain yang masuk. *Feedback* berupa komplain dari pelanggan ini dapat digabungkan menjadi laporan mingguan maupun bulanan yang dapat dengan cepat diserap oleh perusahaan dan dapat mengetahui sejumlah masalah yang ada di perusahaan. Selain itu, banyak keuntungan yang dapat diperoleh dari penanganan komplain yang efisien, salah satunya dapat menyusun rencana strategis perusahaan (Mitchell, 1993). Tujuan dari menganalisis komplain agregat antara lain mengidentifikasi penyebab komplain dan dapat memodifikasi proses bisnis untuk mengeliminasi komplain di masa yang akan datang (Schibrowsky dan Lapidus, 1994).

### 2.2.2. Text Mining

Sebagian besar data perusahaan disimpan dalam bentuk teks, beberapa peneliti menyebutkan bahwa persentase data dalam bentuk teks mencapai 85 hingga 90% (Mason-Jones dan Towill, 1999). Data inilah yang kemudian akan digunakan untuk mengembangkan kegiatan bisnis suatu perusahaan. *Text mining* dapat digunakan untuk menemukan pola tersembunyi dari suatu teks yang tidak terstruktur (Natarajan, 2005). Sebelum data dalam teks dapat disimpulkan dan dipetakan, terdapat beberapa tahapan dalam *text mining* yang harus dilalui, yaitu: *text preprocessing*, *text transformation*, dan *feature selection / attribute selection*, *text mining methods*, interpretasi atau evaluasi (Patel dan Soni, 2012). Proses umum *text mining* dapat dilihat pada Gambar 2.2.



Gambar 2.2. Proses Umum *Text Mining* (Patel dan Soni, 2012)

Tahapan awal pada *text mining* adalah *text preprocessing*. Adapun *text preprocessing* diawali dengan pemecahan dokumen menjadi kalimat, mengubah huruf besar dalam dokumen tersebut menjadi huruf kecil, *stopword removal*, *stemming*, dan diakhiri dengan *n-gramming* jika diperlukan (Veritawati dkk, 2015). Kesulitan dalam pengolahan data tekstual yang tidak terstruktur seringkali menyebabkan perusahaan kurang memanfaatkan data tersebut. Padahal data tersebut bisa digunakan untuk penjaminan kualitas maupun atau untuk memantau kinerja perusahaan (Ashton dkk, 2014). *Text mining* digunakan untuk mempermudah dan mempercepat proses pemanfaatan data tersebut tanpa harus mengeluarkan banyak sumber daya.

### **2.2.3. One Sample Sign Test**

*One Sample Sign Test* merupakan salah satu uji non-parametrik. *One Sample Sign Test* digunakan untuk menguji hipotesis dengan melihat median pada satu populasi (Corrado dan Zivney, 1992). *One Sample Sign Test* biasanya digunakan apabila data diambil tidak memenuhi asumsi normalitas (Shier, 2004). Pada penelitian ini, *One Sample Sign Test* digunakan untuk membandingkan pendapat staf dan supervisor *Quality Assurance* yang digunakan untuk mengelola komplain sebelum dan sesudah menggunakan *text mining*. Uji ini digunakan untuk memastikan apakah penggunaan *text mining* dapat mempermudah staf *Quality Assurance* dalam mengelola *e-mail* komplain, mempersingkat waktu yang diperlukan untuk mengelola *e-mail* komplain, dan apakah hasil pengelolaan *e-mail* komplain yang diperoleh sesuai dengan kenyataan.