

# Implementasi Pohon Keputusan pada Analisis Sentimen dengan Pendekatan Leksikal terhadap Ulasan Produk di *Amazon.com*

Reyhan Naufal Hakim/13517029  
Program Studi Teknik Informatika  
Sekolah Teknik Elektro dan Informatika  
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia  
reyhan\_kim@students.itb.ac.id

**Abstrak**—Seiring dengan perkembangan adopsi internet di dunia, industri *e-commerce* juga semakin populer, contohnya Amazon. Selain *e-commerce* dinilai praktis, pembeli juga dapat secara langsung memberikan ulasan terhadap produk yang dibelinya dalam bentuk tulisan. Analisis Sentimen adalah proses untuk menentukan apakah opini seseorang itu tergolong positif, negatif, atau netral. Walaupun di Amazon sudah terdapat sistem *5-star rating*, seringkali bintang yang pengulas berikan tidak sesuai dengan ulasan tertulis yang diberikannya. Analisis sentimen ini dilakukan dengan menerapkan konsep pohon keputusan. Pada makalah ini, penulis akan melakukan analisis sentimen terhadap ulasan produk dengan data sampel berkategori *Digital Video Games* yang ditulis pada tahun 2006 hingga 2015. Hasil dari analisis sentimen ini kemudian akan dibahas pemanfaatannya dalam dunia bisnis.

**Kata Kunci**—analisis sentimen, amazon, pohon keputusan, ulasan produk

## I. PENDAHULUAN

Analisis sentimen atau *opinion mining* adalah bidang keilmuan yang menganalisis opini manusia, perasaan (sentimen), evaluasi, penilaian, sikap, dan emosi terhadap suatu entitas seperti produk, jasa, organisasi, individu, isu, kegiatan, topik, dan fitur-fitur terkait entitas tersebut. Analisis sentimen dilakukan untuk melihat kecenderungan opini manusia terhadap entitas tersebut, apakah cenderung berpandangan positif, negatif, atau netral. Dengan banyaknya data yang kita miliki di internet terkait ulasan produk, forum diskusi, blog, Twitter, dan jejaring sosial lainnya, analisis sentimen dapat membantu manusia dalam mengambil keputusan terkait strategi bisnis dan pengembangan produk. Tidak hanya untuk kepentingan bisnis, analisis sentimen juga dapat digunakan untuk menentukan pola perilaku masyarakat mengenai isu-isu hangat yang sedang terjadi di media daring.

Amazon adalah salah satu *marketplace* daring terbesar di dunia, dengan jumlah kunjungan sebanyak 2,84 miliar<sup>6</sup> pada bulan November 2018. Setiap produk yang dijual di Amazon memiliki kolom ulasan, di mana pembeli dapat memberikan pendapatnya mengenai kualitas produk yang telah dibelinya. Aspek yang dinilai dapat berupa kualitas produk ataupun penanganan penjual terhadap produk tersebut sebelum pengiriman. Selain ulasan tertulis, pembeli juga dapat memberikan penilaian berupa *rating* yang berbentuk bintang berskala 1 sampai 5. Namun, menurut Bo Pang dan Lillian Lee dalam buku “*Opinion Mining and Sentiment Analysis*”, terdapat

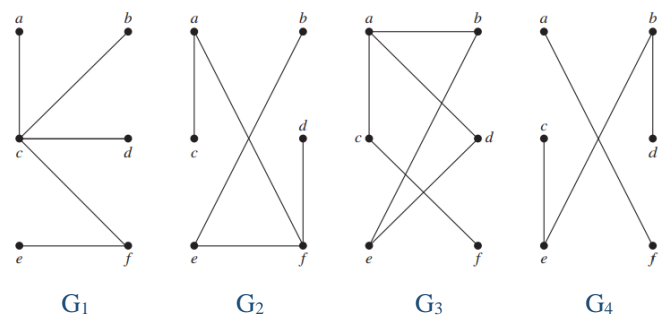
indikasi bahwa *rating* pengguna bisa saja bias dan membutuhkan koreksi, dan *classifier* pada analisis sentimen dapat membantu dalam mengoreksi bias pada *rating* pengguna. Salah satu skenario yang menyebabkan bias pada *rating* pengguna adalah saat pengguna secara tidak sengaja memberikan *rating* rendah untuk ulasan tertulis yang positif.

Dalam *Registry of Open Data on AWS*, Amazon menyediakan data ulasan produk sebelum tahun 2016 untuk keperluan riset *Natural Language Processing*, *Information Retrieval*, dan *Machine Learning*. Data *realtime* hanya bisa didapatkan oleh penjual di Amazon sehingga penulis hanya akan menggunakan *dataset* yang ada di *registry* sebagai contoh penerapan analisis sentimen untuk keperluan bisnis dan pengembangan produk.

## II. LANDASAN TEORI

### A. Pohon

Pohon pertama kali digunakan pada tahun 1857, ketika seorang matematikawan Inggris bernama Arthur Cayley menggunakan pohon untuk menghitung jumlah senyawa kimia tertentu. Pohon adalah graf yang khusus, yaitu graf tak-berarah terhubung yang tidak mengandung sirkuit.



**Gambar 2.1 Graf**

Sumber: *Discrete Mathematics and Its Applications*, Kenneth H. Rosen, Halaman 746

Dengan definisi pohon yang telah disebutkan, pada gambar 2.1, graf  $G_1$  dan  $G_2$  adalah pohon karena keduanya merupakan graf terhubung tanpa sirkuit.  $G_3$  bukan sebuah pohon karena memiliki sirkuit  $e, b, a, d, e$ .  $G_4$  bukan sebuah pohon karena  $G_4$  bukan graf terhubung.

Menurut Rinaldi Munir (2016), pohon memiliki sifat-sifat (*properties*). Misalkan  $G = (V, E)$  adalah graf tak-berarah

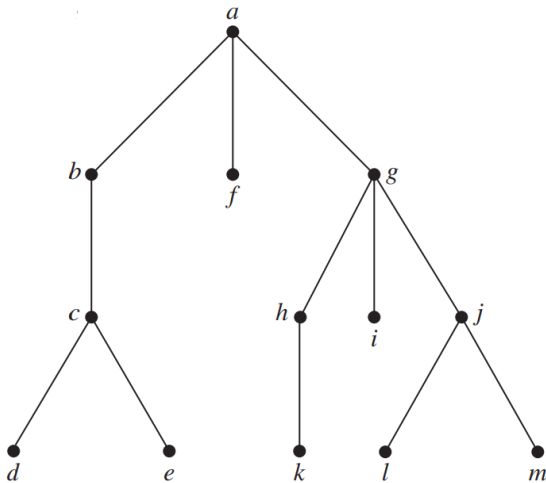
sederhana dan memiliki jumlah simpul  $n$ . Maka, semua pernyataan di bawah ini adalah ekuivalen:

1.  $G$  adalah pohon.
2. Setiap pasang simpul di dalam  $G$  terhubung dengan lintasan tunggal.
3.  $G$  terhubung dan memiliki  $m = n - 1$  buah sisi.
4.  $G$  tidak mengandung sirkuit dan memiliki  $m = n - 1$  buah sisi.
5.  $G$  tidak mengandung sirkuit dan penambahan satu sisi pada graf akan membuat hanya satu sirkuit.
6.  $G$  terhubung dan semua sisinya adalah jembatan (jembatan adalah sisi yang bila dihapus menyebabkan graf terpecah menjadi dua komponen).

Semua butir pernyataan di atas juga dapat dianggap sebagai definisi lain dari pohon.

### B. Pohon Berakar

Dalam implementasi pohon, terdapat simpul yang disebut sebagai akar. Melalui akar yang telah ditetapkan, simpul-simpul lainnya dapat dicapai dengan memberi arah pada sisi-sisi pohon yang mengikutinya. Dengan demikian, pohon berakar adalah pohon yang sebuah simpulnya diperlakukan sebagai akar dan sisi-sisinya diberi arah menjauh dari akar tersebut.



**Gambar 2.2 Pohon Berakar**

Sumber: *Discrete Mathematics and Its Applications*, Kenneth H. Rosen, Halaman 748

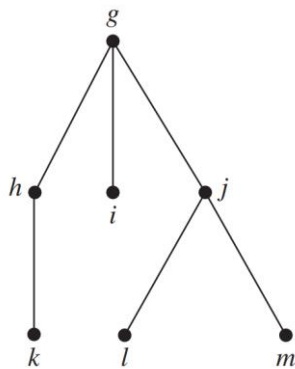
Sebagai contoh, pada gambar 2.2, simpul  $a$  merupakan akar dari sebuah graf. Sesuai dengan definisi graf berakar, simpul lainnya ( $b, c, d, e, f, g, h, k, l, j, l, m$ ) dapat dicapai dari simpul  $a$  yang sudah ditetapkan sebelumnya sebagai akar.

Terdapat beberapa terminologi terkait pohon berakar, di antaranya:

1. Akar  
Akar adalah simpul yang mempunyai derajat-masuk sama dengan nol. Dengan kata lain, akar hanya memiliki sisi keluar dan tidak memiliki sisi masuk. Pada gambar 2.2, simpul  $a$  merupakan contoh dari akar.
2. Daun  
Daun adalah simpul yang mempunyai derajat-keluar

sama dengan nol. Dengan kata lain, daun hanya memiliki sisi masuk dan tidak memiliki sisi keluar. Daun juga sering disebut sebagai simpul terminal. Pada gambar 2.2, simpul  $d, e, f, g, k, l, l,$  dan  $m$  adalah contoh dari daun.

3. Anak (*child* atau *children*)  
Suatu simpul  $y$  dikatakan anak simpul  $x$  bila ada sisi dari simpul  $x$  ke simpul  $y$ . Sebagai contoh, pada gambar 2.2, simpul  $b, f,$  dan  $g$  merupakan anak dari simpul  $a$ .
4. Orangtua (*parent*)  
Suatu simpul  $x$  dikatakan orangtua dari simpul  $y$  bila ada sisi dari simpul  $x$  ke simpul  $y$ . Sebagai contoh, pada gambar 2.2, simpul  $a$  merupakan orangtua dari simpul  $b, f,$  dan  $g$ .
5. Lintasan (*path*) dan Panjang Lintasan  
Lintasan dari simpul  $a_1$  ke  $a_n$  merupakan runtutan simpul-simpul  $a_1, a_2, \dots, a_n$  sedemikian sehingga  $a_i$  adalah orangtua dari  $a_{i+1}$  untuk  $1 \leq i < n$ . Panjang lintasan adalah jumlah sisi yang dilalui dalam suatu lintasan, yaitu  $n - 1$ . Sebagai contoh, pada gambar 2.2, lintasan  $a$  ke  $d$  adalah  $a, b, c, d$  dengan panjang lintasan adalah 3.
6. Keturunan (*descendant*) dan Leluhur (*ancestor*)  
Jika terdapat lintasan dari simpul  $x$  ke simpul  $y$  dalam suatu pohon, simpul  $x$  adalah leluhur dari simpul  $y$  dan simpul  $y$  adalah keturunan dari simpul  $x$ . Sebagai contoh, pada gambar 2.2, simpul  $a$  merupakan leluhur dari simpul  $h$  dan simpul  $h$  merupakan keturunan dari simpul  $a$ .
7. Saudara Kandung (*sibling*)  
Simpul yang memiliki orangtua yang sama adalah saudara kandung satu sama lain. Sebagai contoh, pada gambar 2.2, simpul  $l$  dan  $m$  merupakan saudara kandung karena simpul  $j$  merupakan orangtua dari simpul  $l$  dan simpul  $m$ .
8. Upapohon (*subtree*)  
Misal  $x$  adalah simpul dalam pohon  $T$ . Upapohon dengan akar  $x$  adalah upagraf  $T' = (V', E')$  sedemikian sehingga  $V'$  mengandung  $x$  dan semua keturunannya dan  $E'$  mengandung sisi-sisi dalam semua lintasan yang berasal dari  $x$ . Sebagai contoh, gambar 2.3 merupakan salah satu upapohon  $T' = (V', E')$  dari pohon pada gambar 2.2, di mana  $V' = \{g, h, k, i, j, l, m\}$  dan  $E' = \{(g, h), (g, i), (g, j), (h, k), (j, l), (j, m)\}$ , dengan  $g$  adalah simpul akarnya.



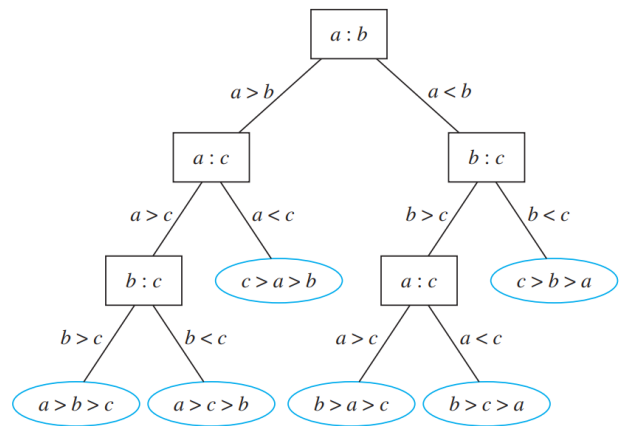
**Gambar 2.3 Sebuah Upapohon**

Sumber: *Discrete Mathematics and Its Applications*, Kenneth H. Rosen, Halaman 748

9. Derajat (*Degree*)  
Derajat sebuah simpul pada pohon berakar adalah jumlah upapohon (atau jumlah anak) pada simpul tersebut. Sebagai contoh, pada gambar 2.2, derajat  $a$  adalah 3 dan derajat  $h$  adalah 1.
10. Simpul Dalam (*Internal Nodes*)  
Simpul dalam adalah simpul yang mempunyai anak. Sebagai contoh, pada gambar 2.2, simpul  $h$  dan  $j$  adalah simpul dalam dan simpul  $i$  bukan merupakan simpul dalam.
11. Aras (*level*) atau Tingkat  
Akar mempunyai tingkat yang sama dengan 0, sedangkan simpul lainnya memiliki tingkat sama dengan  $1 +$  panjang lintasan dari akar ke simpul tersebut. Pada makalah ini, sebagai konvensi, kita memulai penomoran tingkat dari 0. Sebagai contoh, pada gambar 2.2, tingkat dari simpul  $a$  adalah 0 dan tingkat dari simpul  $i$  adalah 2.
12. Tinggi (*height*) atau Kedalaman (*depth*)  
Tinggi maksimum dari suatu pohon disebut sebagai tinggi atau kedalaman pohon tersebut. Sebagai contoh, pada gambar 2.2, tinggi dari pohon tersebut adalah 3.

### C. Pohon Keputusan

Pohon keputusan adalah pohon berakar yang digunakan untuk memodelkan suatu permasalahan dengan serangkaian pengambilan keputusan untuk mencapai suatu solusi. Secara struktur, pohon keputusan adalah pohon berakar yang setiap simpul dalamnya berkorespondensi dengan suatu keputusan (*decision*) dan setiap daun berkorespondensi dengan suatu solusi. Sebagai contoh, pada gambar 2.4, diberikan sebuah pohon keputusan untuk mengurutkan 3 buah elemen berbeda.



**Gambar 2.4 Pohon Keputusan untuk Mengurutkan 3 buah Elemen Berbeda**

Sumber: *Discrete Mathematics and Its Applications*, Kenneth H. Rosen, Halaman 761

### D. Analisis Sentimen

Analisis sentimen atau *opinion mining* adalah bidang keilmuan yang menganalisis opini manusia, perasaan (sentimen), evaluasi, penilaian, sikap, dan emosi terhadap suatu entitas seperti produk, jasa, organisasi, individu, isu, kegiatan, topik, dan fitur-fitur terkait entitas tersebut. Riset dan penelitian mengenai analisis sentimen diprakarsai oleh Bo Pang pada tahun 2002. Menurut Anaïs Collomb, secara teknis terdapat beberapa cara untuk melakukan analisis sentimen: pembelajaran mesin, *lexicon-based*, pendekatan statistik, dan pendekatan *rule-based*.

1. Metode pembelajaran mesin menggunakan algoritma untuk menentukan sentimen manusia dengan mempelajari dataset yang sudah diketahui.
2. Pendekatan *lexicon-based* akan menghitung polaritas sentimen untuk suatu teks dengan menggunakan orientasi semantik dari kata-kata atau kalimat-kalimat yang ada pada teks tersebut.
3. Pendekatan *rule-based* akan mencari kata-kata opini dalam teks dan mengklasifikasi kata tersebut berdasarkan jumlah kata-kata positif dan kata-kata negatif yang ada.
4. Model statistik merepresentasikan setiap teks sebagai kombinasi antara berbagai aspek beserta penilaian. Aspek dan penilaian yang didapatkan dari teks akan diasumsikan dapat diubah menjadi distribusi multinomial. *Classifier* lalu mengklasterisasi istilah-istilah kunci menjadi aspek-aspek dan sentimen-sentimen yang ada pada teks menjadi penilaian-penilaian (*ratings*).

Pada makalah ini, penulis akan menggunakan pendekatan *lexicon-based* dalam melakukan analisis sentimen. Pendekatan *lexicon-based* dipilih karena penulis ingin menentukan kualitas dari produk yang dijual dengan mempertimbangkan ulasan teks

yang diberikan pengguna sebagai pembanding terhadap *rating* pengguna.

### III. PENERAPAN POHON KEPUTUSAN UNTUK ANALISIS SENTIMEN TERHADAP ULASAN PRODUK

#### A. Persiapan Data (*Data Cleaning and Data Preparation*)

Sebelum melakukan analisis sentimen terhadap data, kita perlu mengolah data kita terlebih dahulu karena data teks yang kita dapatkan mungkin saja berisi karakter atau kata-kata yang tidak relevan. Karakter atau kata-kata yang tidak relevan ini perlu diubah atau bahkan dihapus sehingga kita hanya menganalisis kata-kata yang relevan pada analisis sentimen nanti. Menurut Matthew Mayo, terdapat 3 proses utama dalam mempersiapkan data: *Tokenization*, *Normalization*, dan *Noise Removal*. Pada makalah ini, penulis hanya melakukan *noise removal* karena mayoritas data tidak memerlukan tokenisasi lagi dan normalisasi dikhawatirkan menghilangkan konteks dari ulasan pembeli. Selain *noise removal*, penulis akan menggunakan dataset dari Bing Liu untuk menentukan leksikon yang bernilai positif dan leksikon yang bernilai negatif.

##### 1. Noise Removal

Sebagai contoh, perhatikan data berikut:

Really thought it would be next to impossible to top FO3.... but boy did Obsidian come through!! Have played all content including DLC's (for those looking to purchase this, please note the packs are NOT content per say, they give you items at the beginning of the game, and the Gun Runners Arsenal adds certain weapons and weapon-mods to the game).  
Awesome barely covers it, plus (if you're into it) the modding is awesome, effectively dozens of extra DLC's if you want to go through them all!!

Pada data di atas, terdapat *noise* berupa kode ASCII '#' yang menunjukkan tanda petik dua (") dan *line break* '\n'. Kedua karakter ini tidak diperlukan dalam proses analisis sentimen dan dapat mempengaruhi akurasi analisis sentimen sehingga harus dihapus. Terlihat bahwa '\n' tidak hanya dihapus, tetapi digantikan oleh karakter spasi. Penulis melakukan ini agar struktur pemisah antara kata yang satu dengan kata yang lain tetap ada walaupun *line break* dihapus.

Hasil *noise removal*:

Really thought it would be next to impossible to top FO3.... but boy did Obsidian come through!! Have played all content including DLC's (for those looking to purchase this, please note the packs are NOT content per say, they give you items at the beginning of the game, and the Gun Runners Arsenal adds certain weapons and weapon-mods to the game). Awesome barely covers it, plus (if you're into it) the modding is awesome, effectively dozens of extra DLC's if you want to go through them all!!

##### 2. Menentukan Leksikon Bernilai Positif dan Leksikon Bernilai Negatif

Untuk dapat menentukan polaritas dari ulasan pembeli, penulis menentukan nilai masing-masing kata yang bernilai positif dan negatif. Setiap kata positif akan bernilai 1 dan setiap kata negatif bernilai -1. Kata-kata yang tidak diberi nilai positif maupun negatif disebut sebagai kata netral dengan nilai 0. Berikut adalah tabel yang menunjukkan beberapa kata positif dan negatif yang sudah diberi label:

Kata	Nilai
<i>top</i>	1
<i>awesome</i>	1
<i>effectively</i>	1
<i>smooth</i>	1
<i>good</i>	1
<i>goood</i>	1
<i>junk</i>	-1
<i>blind</i>	-1
<i>trash</i>	-1
<i>sloow</i>	-1
<i>slow</i>	-1
<i>sloww</i>	-1

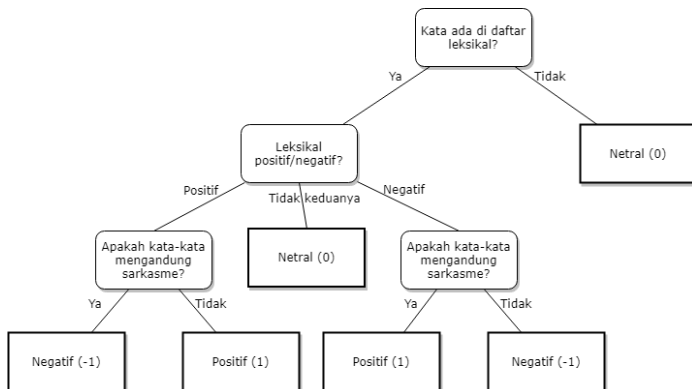
Terlihat bahwa terdapat kata-kata yang tidak sesuai dengan ejaannya. Menurut Bing Liu, ini sebaiknya dilakukan karena pada kenyataannya kata-kata yang salah dari segi ejaan ini sering dipakai orang untuk mengungkapkan sentimen di internet.

##### B. Decision Tree (DT) Classifier

*Decision Tree Classifier* akan mengklasifikasikan kata-kata untuk mencapai suatu kepolaritasan. Kepolaritasan ini direpresentasikan oleh daun pada pohon keputusan, yang dapat juga disebut sebagai solusi polaritas untuk setiap teks ulasan. Kepolaritasan dapat berupa polaritas positif, netral, dan negatif.

*Decision Tree Classifier* dapat digunakan tanpa perlu kita implementasikan sendiri dengan memanfaatkan pustaka `sklearn.tree.DecisionTreeClassifier` pada

python. Karena *Decision Tree Classifier* merupakan suatu metode klasifikasi untuk *supervised learning*, kita dapat dengan mudah menentukan apa yang sebenarnya terjadi di balik algoritma *Decision Tree Classifier* dalam bentuk pohon.



**Gambar 2.5 Pohon Keputusan untuk Setiap Kata pada Teks Ulasan**

Pohon keputusan pada gambar 2.5 berlaku untuk setiap kata, termasuk kata yang tidak menunjukkan sentimen. Kata-kata yang tidak menunjukkan sentimen akan digolongkan sebagai kata netral yang bernilai nol. Deteksi sarkasme dapat dilakukan dengan membandingkan *rating* bintang yang diberikan pengguna dengan teks ulasan yang menjadi objek analisis sentimen. Jika kata-kata pada teks ulasan secara leksikal bermakna positif tetapi *rating* yang diberikan rendah, maka penulis mengasumsikan ulasan tersebut adalah pernyataan yang bersifat sarkasme.

Dengan menurunkan teorema Bayes', yang menyatakan bahwa probabilitas kondisional dapat dirumuskan sebagai

$$p(C_k | \mathbf{x}) = \frac{p(C_k) p(\mathbf{x} | C_k)}{p(\mathbf{x})}$$

untuk setiap hasil K atau class  $C_k$ , nilai polaritas setiap ulasan dapat dikalkulasikan dengan persamaan berikut

$$Polaritas = \frac{(n_a \times w_a) + (n_b \times w_b)}{n}$$

Dengan  $n_a$  frekuensi kejadian a,  $n_b$  frekuensi kejadian b,  $w_a$  nilai dari kata a, dan  $w_b$  nilai dari kata b, dan n adalah frekuensi kata secara total dalam sebuah ulasan.

### C. Contoh Analisis

#### 1. Hasil Analisis Positif

Ulasan:

My granddaughter loves this game. She plays it all the time; I have to make her stop.

Analisis leksikal:

My	0
Granddaughter	0
Loves	1
This	0
Game	0
She	0
Plays	0
It	0
All	0
The	0
Time	0
I	0
Have	0
To	0
Make	0
Her	0
Stop	0

Dengan demikian, polaritas dari contoh ulasan di atas adalah  $((1 \times 1) + (0 \times -1)) / 17 = 0.059$ . Karena nilai polaritas  $> 0$ , ulasan ini termasuk ulasan positif.

#### 2. Hasil Analisis Negatif

Ulasan:

This is just plain junk. When you get to heads up, the player always folds when you are big blind. It will take you hours to win. Nothing but trash.

Analisis leksikal:

This	0
Is	0
Just	0
Plain	0
Junk	-1
When	0
You	0
Get	0
To	0
Heads	0
Up	0
The	0
Player	0
Always	0
Folds	0
When	0
You	0
Hours	0
To	0
Win	0
Nothing	0
But	0
trash	-1

Dengan demikian, polaritas dari contoh ulasan di atas adalah  $((0 \times 1) + (2 \times -1)) / 23 = -0.087$ . Karena nilai polaritas  $< 0$ , ulasan ini termasuk ulasan negatif.

#### IV. APLIKASI ANALISIS SENTIMEN UNTUK KEPERLUAN BISNIS

Di dunia bisnis, analisis sentimen semakin gencar diadopsi. Istilah yang sangat dikenal sekarang adalah *business intelligence*, di mana praktisi bisnis (*entrepreneurs*) memanfaatkan teknologi analisis data untuk keperluan strategi bisnis dan pengembangan produk. Analisis sentimen sendiri adalah bagian dari *business intelligence*.

##### A. Mengetahui Fitur Produk yang Disukai atau Tidak Disukai Pembeli

Dengan menganalisis target dari sentimen pelanggan, pelaku bisnis dapat mengetahui fitur apa yang disukai atau tidak disukai pelanggan dari produknya. Misal,

One of the best. Love the story, graphics, sound, specially the dancing part. Really recommend this to people that like the challenge and have fun.

Everyone knows the problems with Destiny. But I didn't really get upset about it until I paid for the Expansions and realize they're basically all of the same enemies, except with some newer weapons and higher level gear. So you're basically paying to upgrade gear. Why didn't they just go with free-to-play model, then? Because they're Bungie/Activision and they can sell you whatever they want.

Kedua ulasan di atas merupakan sampel ulasan dari produk *Destiny - Expansion Pass - PlayStation 4* yang dijual di Amazon.com. Dari ulasan pertama, kita dapat menyimpulkan bahwa hal yang disukai pemain dari permainan ini adalah alur ceritanya, grafiknya, dan efek suaranya. Akan tetapi, melalui ulasan kedua, kita dapat mengetahui bahwa pemain tidak menyukai konsep ekspansi berbayar yang dinilai kurang memberi konten lebih kepada pemain. Jika data analisis sentimen ini diketahui oleh pengembang game, pengembang game dapat memberikan nilai lebih (*added value*) kepada pemain dengan mempertahankan karakteristik produk yang disukai pemain dan membenahi karakteristik produk yang tidak disukai pemain.

##### B. Mengetahui Mispersepsi atau Masalah yang Dialami Pengguna

Terkadang, terdapat mispersepsi antara pihak pengembang dengan pengguna. Misal, konten dari permainan dirasa kurang menarik, padahal pengembang sudah

menyediakan pembaharuan konten setiap pekan dan hanya pengguna saja yang belum memperbaharui perangkat lunaknya. Kasus lain adalah di mana terdapat *bug* yang membuat pengguna tidak dapat menikmati produk yang dibelinya, seperti:

It never worked! It still does not work. Amazon reimbursed me. EA customer support was atrocious. Had me do stuff all day. And then ended up being a bug in their software that has not yet been scheduled to fix.

Dengan menggunakan analisis sentimen, masalah ini dapat ditangani dengan lebih cepat. Jika pengembang sadar akan masalah ini, mereka dapat segera melakukan evelasi terhadap *issue* yang disampaikan pemain.

#### V. KESIMPULAN

Pemanfaatan pohon keputusan dalam analisis sentimen dapat menyelesaikan banyak masalah terkait kebutuhan bisnis dan pengembangan produk dengan cepat. Pemanfaatan pohon keputusan tentu dapat diaplikasikan lebih luas lagi, seperti untuk analisis kondisi politik, sosial budaya, bahkan analisis trend di sosial media.

#### VI. UCAPAN TERIMA KASIH

Penulis berterima kasih kepada Dr. Ir. Rinaldi Munir yang telah memberikan pembelajaran dan bimbingan dalam kuliah Matematika Diskrit. Penulis juga mengucapkan terima kasih kepada seluruh pihak yang telah memberikan dukungan baik secara fisik maupun moral sehingga penyusunan makalah ini dapat selesai tepat waktu.

#### DAFTAR PUSTAKA

- [1] B. Pang, L. Lee, *Opinion Mining and Sentiment Analysis*, Sunnyvale, CA: NOW, 2008, pp. 1-9.
- [2] B. Pang, L. Lee, *A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts*, Ithaca, NY: Cornell University, 2004.
- [3] X. Fang, J. Zhan, *Sentiment Analysis Using Product Review Data*. Greensboro, NC: Springer, 2015.
- [4] M. Hu, B. Liu, "Mining and Summarizing Customer Reviews". Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), Aug 22-25, 2004, Seattle, Washington, USA.
- [5] R. Munir, *Matematika Diskrit*, Revisi keenam, Bandung, Indonesia: Penerbit Informatika, 2016.
- [6] Rosen, Kenneth H., *Discrete Mathematics and Its Applications*, NY: McGraw-Hill International, 2012, 7<sup>th</sup> ed.
- [7] <https://www.similarweb.com/website/amazon.com>, diakses pada 9 Desember 2018, 10.41 WIB.
- [8] <https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html>, diakses pada 9 Desember 2018, 21.33 WIB

## PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 10 Desember 2017

Reyhan Naufal Hakim  
13517029