

BAB I

PENDAHULUAN

1.1 Latar Belakang

Ketertarikan masyarakat untuk berwirausaha dan berbisnis saat ini menunjukkan angka yang menjanjikan bagi Indonesia. Dengan bermunculannya bisnis-bisnis baru yang menghasilkan produk barang jadi dan jasa yang berkualitas dan mengkonsumsi produk dalam negeri tentunya akan berdampak positif pada negara yaitu mengembalikan perekonomian negara yang sehat. Bahkan, beberapa produk dalam negeri yang berkualitas telah dijual-belikan dan dikonsumsi oleh beberapa negara asing. Di bawah ini adalah data yang menunjukkan peningkatan UMKM (Usaha Mikro Kecil dan Menengah) dari tahun 1997 sampai tahun 2012 yang dicatat oleh BPS (Badan Pusat Statistik).

Tabel Perkembangan UMKM pada Periode 1997 -2012

No.	Indikator	Satuan	1997	1998	1999	2000	2001	2002	2003	2004
1	Jumlah UMKM	Unit	39 765 110	36 813 578	37 911 723	39 784 036	39 964 080	41 944 494	43 460 242	44 777 387
2	Pertumbuhan Jumlah UMKM	Persen		-7,42	2,98	4,94	0,45	4,96	3,61	3,03
3	Jumlah Tenaga Kerja UMKM	Orang	65 601 591	64 313 573	67 169 844	72 704 416	74 687 428	77 807 897	81 942 353	80 446 600
4	Pertumbuhan Jumlah Tenaga Kerja UMKM	Persen		-1,96	4,44	8,24	2,73	4,18	5,31	-1,83
5	Sumbangan PDB UMKM (harga konstan)	Rp. Miliar	363 200,440	552 945,40	647 475,960	760 089,450	791 597,40	829 616,40	876 123,40	924 483,60
6	Pertumbuhan sumbangan PDB UMKM	Persen		52,24	17,10	17,39	4,15	4,80	5,61	5,52
7	Nilai Ekspor UMKM	Rp. Miliar	39 277,070	69 315,40	52 594,120	75 448,610	80 846,520	87 290,030	77 096,710	95 548,240
8	Pertumbuhan Nilai Ekspor UMKM	Persen		76,48	-24,12	43,45	7,15	7,97	-11,68	23,93

No.	Indikator	Satuan	2005	2006	2007	2008	2009	2010	2011	2012
1	Jumlah UMKM	Unit	47 017 062	49 021 803	50 145 800	51 409 612	52 764 603	53 823 732	55 206 444	56 534 592
2	Pertumbuhan Jumlah UMKM	Persen		5,00	4,26	2,29	2,52	2,64	2,01	2,57
3	Jumlah Tenaga Kerja UMKM	Orang	83 586 616	87 909 598	90 491 930	94 024 278	96 211 332	99 401 775	101 722 458	107 657 509
4	Pertumbuhan Jumlah Tenaga Kerja UMKM	Persen		3,90	5,17	2,94	3,90	2,33	3,32	2,33
5	Sumbangan PDB UMKM (harga konstan)	Rp. Miliar	979 712,50	1 032 573,90	1 099 301,10	1 165 753,20	1 212 599,30	1 282 571,80	1 369 326,00	1 504 928,20
6	Pertumbuhan sumbangan PDB UMKM	Persen		5,97	5,40	6,46	6,04	4,02	5,77	6,76
7	Nilai Ekspor UMKM	Rp. Miliar	110 338,06	123 767,87	140 363,84	178 008,28	162 254,52	175 894,89	187 441,82	208 067,00
8	Pertumbuhan Nilai Ekspor UMKM	Persen		15,48	12,17	13,41	26,82	-8,85	8,41	6,56

- Keterangan:
1. Data UMKM pada periode pemerintahan Orde Lama tidak tersedia. Data UMKM pada periode Orde Baru mulai tersedia pada tahun 1997.
 2. Sumbangan PDB UMKM pada tahun 1997-2000 masih didasarkan atas harga berlaku. Sumbangan PDB UMKM pada tahun 2001-2012 didasarkan atas harga konstan tahun 2000.
 3. Sumbangan PDB mencakup migas dan non migas.
 4. Perhitungan sumbangan PDB UMKM pada tahun 2012 masih sangat sementara dan saat ini masih dihitung ulang.
 5. Nilai Ekspor UMKM hanya didasarkan pada data di sektor pertanian (pertanian, perikanan, kelautan, peternakan, kehutanan, perkebunan), industri pengolahan, dan pertambangan/pengalihan.

Gambar 1.1 Perkembangan UMKM periode 1997 - 2012 (Sumber: <http://www.bps.go.id/>)

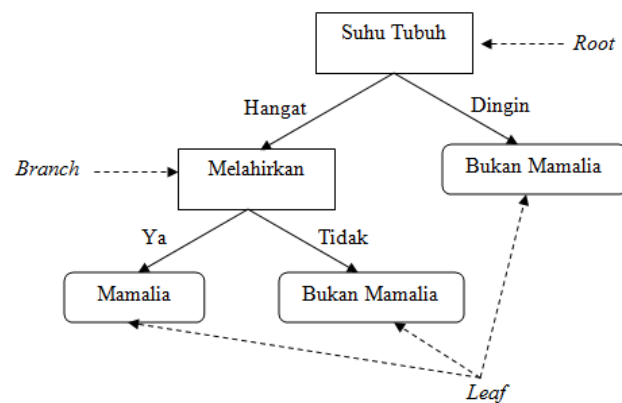
Berkembangnya dunia bisnis yang beriringan dengan kemajuan teknologi informasi melahirkan gaya baru berbisnis secara *online*. Dengan berbisnis secara *online*, aktivitas bisnis akan lebih efisien dan efektif karena dapat mencapai jangkauan yang lebih luas dengan waktu yang lebih singkat dan tenaga yang lebih sedikit misalnya berkomunikasi dengan pelanggan melalui telepon, melakukan pemasaran melalui media *internet* atau *televisi*, melakukan transaksi melalui ATM (*Automatic Teller Machine*), dan sebagainya. Seorang ahli berpendapat pemasaran melalui media *internet* akan lebih efektif dibandingkan dengan media tradisional seperti surat kabar dan reklame. Hal ini dipercaya karena kebiasaan masyarakat saat ini lebih sering menggunakan waktu senggang mereka di komputer dan perangkat *mobile*. Selain itu, beliau menambahkan bahwa pemasaran melalui media *internet* lebih mudah untuk dianalisa karena jumlah pengunjung yang mengunjungi halaman situs mereka di *internet* dapat dilacak (Liputan 6, 2014).

Beberapa pelaku bisnis saat ini memanfaatkan media sosial yang populer sebagai media pemasaran mereka. Contohnya seperti Lazada, Zalora, Traveloka, dan sebagainya pada *facebook*. Bahkan *twitter* pun yang dirancang sebagai situs mikroblog menjadi sasaran beberapa pelaku bisnis seperti myCampus Store ITB, Maicih, dan sebagainya. Berbeda dengan *facebook*, *twitter* memberikan penggunaan yang lebih simpel dengan informasi yang lebih singkat dan terbatas karena *twitter* merupakan situs mikroblog. Mikroblog berarti cara berbagi informasi dengan singkat dan tidak memakan waktu yang lama (Ebner Martin, 2008). Keterbatasan informasi pada *twitter* juga ada pada profile pengguna karena di *twitter* tidak ada informasi mengenai jenis kelamin dan usia dari pengguna.

Keterbatasan pada *twitter* tersebut tentu akan mempersulit proses analisa pasar karena di dalam bisnis mensegmentasikan pasar dibutuhkan untuk menentukan strategi pemasaran yang tepat (R. Smith W, 1995) (Nurbaity A, 2004). Dengan demikian, pelaku bisnis *online* yang menggunakan media sosial *twitter* tidak dapat mensegmentasikan pasar bisnisnya secara langsung. Berpegang pada teori psikolinguistik, karakteristik seseorang dapat diketahui dari tata bahasanya. Psikolinguistik ini adalah gabungan ilmu psikologi dan ilmu bahasa yang mempelajari tentang keterkaitan antara karakteristik seseorang dengan tata bahasa yang mereka gunakan (Diana L, 2013) (Wahyono H, 2012). Dengan begitu, permasalahan tentang keterbatasan informasi pada *twitter* dapat diselesaikan dengan mempraktekan ilmu psikolinguistik.

Meskipun begitu, mensegmentasikan pasar bisnis *online* yang telah memiliki banyak pelanggan dan dilakukan pada media sosial *twitter* akan memakan waktu yang cukup lama. Waktu yang digunakan untuk mensegmentasikan pasar tersebut kurang efisien. Secara teori, permasalahan efisiensi waktu ini dapat diselesaikan dengan menggunakan *machine learning*. *Machine learning* adalah cabang dari ilmu komputer yang digunakan untuk menggali dan mempelajari pola dari data-data berskala besar (Alpaydin Ethem, 2004:2). Contohnya adalah data pada supermarket terdapat beberapa pola seperti produk terlaris pada area tersebut, pembelian produk roti dan mentega dalam satu transaksi, dan sebagainya. Hal ini dilakukan untuk mendapatkan informasi tersembunyi yang dapat digunakan setelahnya misalnya memposisikan produk-produk yang terkait lebih dekat, memperbanyak stok dan varian produk yang sangat laris pada area tersebut.

Machine learning menggunakan algoritma pemrograman untuk menggali data berskala besar tersebut. Salah satu algoritma yang dapat diaplikasikan pada *machine learning* adalah *Decision Tree C4.5*. *Decision Tree C4.5* adalah algoritma yang digunakan untuk mengklasifikasikan data dan digambarkan dengan diagram pohon keputusan (Quinlan J.R, 1993:5). Sebagai contoh, di bawah ini adalah contoh diagram pohon sederhana untuk mengklasifikasikan mamalia.



Gambar 1.2 Contoh diagram pohon sederhana untuk mengklasifikasikan mamalia

Berpedoman pada teori-teori tersebut, penulis tertarik untuk melakukan penelitian lebih lanjut mengenai permasalahan ini dan mengangkat penelitian ini menjadi skripsi dengan judul “Sistem Penentu Demografi Pelanggan Berdasarkan Data Tweet Dengan Menggunakan Algoritma Decision Tree C4.5”.

1.2 Perumusan Masalah

Berdasarkan latar belakang, timbul perumusan masalah sebagai berikut :

1. Pola seperti apa yang dapat digunakan untuk mengekstrak informasi yang dibutuhkan dari data tweet agar dapat menghasilkan demografi pelanggan?
2. Bagaimana pola tersebut bisa digunakan untuk mengekstrak demografi pelanggan?

1.3 Batasan Masalah

Agar penelitian ini terarah, batasan masalah untuk penelitian ini yaitu :

1. Tweet yang diproses adalah tweet berbahasa Indonesia.
2. Pengguna yang diproses adalah pengguna aktif yang memiliki minimal 200 tweet.
3. Unsur demografi yang digunakan adalah unsur demografi dengan sudut pandang ekonomi yang umum (*general*) saja yaitu umur, jenis kelamin, dan status pekerjaan.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah :

1. Mendapatkan pola yang tepat untuk mengekstrak informasi yang dibutuhkan dari data tweet yang kemudian menghasilkan demografi pelanggan.
2. Menggunakan pola tersebut untuk mengekstrak demografi pelanggan.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dalam penelitian ini adalah :

1. Menghasilkan informasi prediksi demografi yang dapat membantu pelaku bisnis dalam pengambilan strategi yang tepat terhadap pasar.
2. Memberikan motivasi untuk penelitian selanjutnya yang berhubungan dengan prediksi demografi.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN

Bab ini menjelaskan mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian.

BAB II TINJAUAN PUSTAKA

Bab ini membahas tentang acuan teori yang digunakan dalam penelitian.

BAB III METODE PENELITIAN

Bab ini membahas tentang penerapan metode kedalam aplikasi yang dibangun.

BAB IV PEMBAHASAN

Bab ini berisikan hasil dari penelitian beserta pembahasannya.

BAB V KESIMPULAN DAN SARAN

Bab ini berisikan kesimpulan dan saran yang diperoleh dari hasil penelitian.