

**ANALISIS BIG DATA PENGGUNAAN INTERNET
PADA PENGGUNA LAYANAN UBIQU**

Tugas Akhir

**Diajukan untuk Memenuhi Salah Satu Persyaratan Mencapai Derajat
Sarjana Komputer**



Dibuat Oleh:

GIVAN ANDRIAN NUGROHO

130707541

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ATMA JAYA YOGYAKARTA
2020**

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

ANALISIS BIG DATA PENGGUNAAN INTERNET PADA PENGGUNA LAYANAN UBIQU

yang disusun oleh

GIVAN ANDRIAN NUGROHO

130707541

dinyatakan telah memenuhi syarat pada tanggal 23 November 2020

		Keterangan
Dosen Pembimbing 1	: Martinus Maslim, ST., MT.	Telah menyetujui
Dosen Pembimbing 2	: Patricia Ardanari, S.Si., M.T	Telah menyetujui
Tim Penguji		
Penguji 1	: Martinus Maslim, ST., MT.	Telah menyetujui
Penguji 2	: Joanna Ardhyanti Mita N, S.Kom., M.Kom	Telah menyetujui
Penguji 3	: Stephanie Pamela Adithama, ST., MT.	Telah menyetujui

Yogyakarta, 23 November 2020

Universitas Atma Jaya Yogyakarta

Fakultas Teknologi Industri

Dekan

ttd

Dr. A. Teguh Siswanto, M.Sc

PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Givan Andrian Nugroho
NPM : 130707541
Program Studi : Informatika
Fakultas : Teknologi Industri
Judul Penelitian : Analisis Big Data Penggunaan Internet Pada
Pengguna Layanan Ubiq

Menyatakan dengan ini:

1. Tugas Akhir ini adalah benar tidak merupakan salinan sebagian atau keseluruhan dari karya penelitian lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta atas penelitian ini, berupa Hak untuk menyimpan, mengelola, mendistribusikan, dan menampilkan hasil penelitian selama tetap mencantumkan nama penulis.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum atas pelanggaran Hak Cipta dalam pembuatan Tugas Akhir ini.

Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 24 Januari 2020

Yang menyatakan,

Givan Andrian Nugroho

130707541

PERNYATAAN PERSETUJUAN DARI INSTANSI ASAL PENELITIAN

Saya yang bertanda tangan di bawah ini:

Nama Lengkap Pembimbing : Gatot Teguh Rahardjo
Jabatan : Manager
Departemen : IT Infrastructure Development

Menyatakan dengan ini:

Nama Lengkap : Givan Andrian Nugroho
NPM : 130707541
Program Studi : Informatika
Fakultas : Teknologi Industri
Judul Penelitian : Analisis Big Data Penggunaan Internet Pada Pengguna Layanan Ubiq
Pengguna Layanan Ubiq

1. Penelitian telah selesai dilaksanakan pada perusahaan.
 2. Perusahaan telah melakukan sidang internal berupa kelayakan penelitian ini dan akan mencantumkan lembar penilaian secara tertutup kepada pihak universitas sebagai bagian dari nilai akhir mahasiswa.
 3. Memberikan kepada Instansi Penelitian dan Universitas Atma Jaya Yogyakarta atas penelitian ini, berupa hak untuk menyimpan, mengelola, mendistribusikan, dan menampilkan hasil penelitian selama tetap mencantumkan nama penulis.
- Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Jakarta, 24 November 2020

Yang menyatakan,

Gatot Teguh Rahardjo
Manager

HALAMAN PERSEMBAHAN

**Teruntuk semua yang masih terus berjuang untuk tetap hidup dikala dunia
yang sedang tidak baik-baik saja,**

Terima kasih.



KATA PENGANTAR

Puji dan syukur penulis sampaikan kepada Tuhan Yang Maha Esa karena berkat rahmat dan karunia-Nya penulis dapat menyelesaikan pembuatan tugas akhir “Analisis Big Data Penggunaan Internet Pada Pengguna Layanan Ubiq ” ini dengan baik.

Penulisan tugas akhir ini bertujuan untuk memenuhi salah satu syarat untuk mencapai derajat sarjana komputer dari Program Studi Informatika, Fakultas Teknologi Industri di Universitas Atma Jaya Yogyakarta.

Penulis menyadari bahwa dalam pembuatan tugas akhir ini penulis telah mendapatkan bantuan, bimbingan, dan dorongan dari banyak pihak. Untuk itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Tuhan Yang Maha Esa yang selalu memberikan hikmat akal budi, dan perlindungan yang diberikan-Nya.
2. Bapak Dr. A. Teguh Siswantoro, selaku Dekan Fakultas Teknologi Industri, Universitas Atma Jaya Yogyakarta.
3. Bapak Martinus Maslim, S.T., MT., selaku Ketua Program Studi Informatika Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta sekaligus Dosen Pembimbing I yang telah membimbing hati dengan tulus mendampingi, membimbing, memberikan masukan yang dapat membangun penulis dari semenjak masuk kedalam Program Studi Informatika hingga saat ini.
4. Ibu Patricia Ardanari, S.Si., M.T, selaku Dosen Pembimbing II yang telah membimbing dan memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.
5. Seluruh Dosen, Staf dan Asisten Dosen Program Studi Informatika Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta yang telah banyak membantu penulis ketika berproses selama masa perkuliahan.
6. Bapak, Mama, Mba Ira, Mba Puji, Dhanty, Elin tercinta yang senantiasa

memberikan dukungan, menyemangati, memfasilitasi dan mendoakan demi kelancaran penulis dalam proses mengerjakan tugas akhir ini.

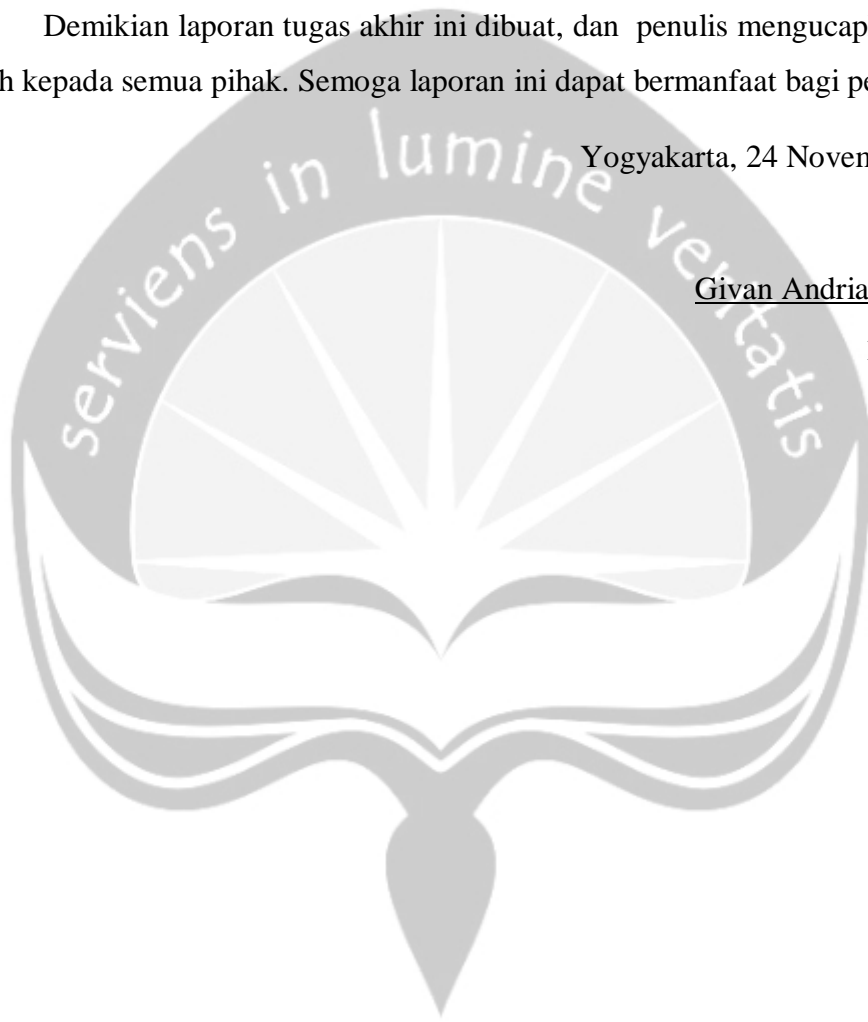
7. Teman-teman Informatika Universitas Atma Jaya Yogyakarta angkatan 2013 atas kebersamaan selama masa perkuliahan.
8. Semua pihak yang tidak dapat penulis sebutkan satu per satu, terima kasih untuk semuanya.

Demikian laporan tugas akhir ini dibuat, dan penulis mengucapkan terima kasih kepada semua pihak. Semoga laporan ini dapat bermanfaat bagi pembaca.

Yogyakarta, 24 November 2020

Givan Andrian Nugroho

130707541



DAFTAR ISI

PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH.....	iii
PERNYATAAN PERSETUJUAN DARI INSTANSI ASAL PENELITIAN	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
INTISARI.....	xiii
BAB I. PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian.....	4
1.5. Metode Penelitian.....	4
1.6. Sistematika Penulisan	5
BAB II. TINJAUAN PUSTAKA	7
2.1 Penelitian Sebelumnya	7
BAB III. LANDASAN TEORI	11
3.1 Big Data.....	11
3.2 Data Mining	12
3.3 Reverse DNS Lookup.....	13
3.4 Pemrograman Python	14
3.5 Ubiqu	14
3.6. Deep Packet Inspection	15
3.7. Ntop.....	16
BAB IV. METODOLOGI PENELITIAN	17
4.1. Objek Penelitian	17
4.2. Jenis Dan Sumber Data.....	17
4.2.1. Jenis Data	17

4.3.	Metode Pengumpulan Data.....	18
4.4.	Tahapan Penelitian Menggunakan Metode Data Mining	18
BAB V. HASIL DAN PEMBAHASAN.....		37
5.1.	Hasil Data Penggunaan Harian	37
5.2.	Hasil Data Penggunaan Berdasarkan Protokol	38
5.3.	Hasil Data Penggunaan Berdasarkan Jenis Paket	40
5.4.	Hasil Data Penggunaan Jenis Produk dan Provinsi.....	41
5.5.	Hasil Data Penggunaan Aplikasi Pengirim Pesan.....	43
5.6.	Hasil Data Penggunaan Situs/Aplikasi Jual Beli Online	44
5.7.	Hasil Data Penggunaan Situs Berita Online	45
5.8.	Hasil Data Penggunaan Situs Perjalanan Online	47
5.9.	Hasil Data Penggunaan Situs Perbankan Online	48
5.10.	Hasil Data Penggunaan Situs Streaming Film Online.....	49
BAB VI. PENUTUP		50
6.1.	Kesimpulan	50
6.2.	Saran.....	50
DAFTAR PUSTAKA		52

DAFTAR GAMBAR

Gambar 4.1 Contoh file .csv yang dihasilkan aplikasi ntop.....	19
Gambar 4.2 Dataset dari pelanggan	20
Gambar 4.3 Script python penggabungan data perhari.....	21
Gambar 4.4 Perbandingan data sebelum dan sesudah script.....	22
Gambar 4.5 Script menjadi satu file	23
Gambar 4.6 Dataset yang berhasil diubah menjadi dataframe	24
Gambar 4.7 Perubahan IP Address Pelanggan	24
Gambar 4.8 Proses rDNS	25
Gambar 4.9 Proses penghitungan request per hari	26
Gambar 4.10 Proses pencarian pelanggan teraktif	27
Gambar 4.11 Proses penghitungan IP tujuan terbanyak	27
Gambar 4.12 Proses pencarian layer 7 protocol teraktif	28
Gambar 4.13 Proses penghitungan pelanggan aktif per produk dan provinsi.....	29
Gambar 4.14 Proses penghitungan provinsi dengan pelanggan teraktif.....	30
Gambar 4.15 Proses penghitungan jumlah request perpaket	31
Gambar 4.16 Proses penghitungan situs film favorit pelanggan.....	32
Gambar 4.17 Proses pencarian situs traveling favorit	33
Gambar 4.18 Proses pencarian aplikasi kirim pesan terfavorit	33
Gambar 4.19 Proses pencarian situs belanja online terfavorit	34
Gambar 4.20 Proses pencarian Internet Banking terfavorit	34
Gambar 4.21 Proses pencarian situs berita terfavorit	35
Grafik 5.1 Pergerakan request perhari	37
Grafik 5.2 Akses berdasarkan protokol perbulan	38
Grafik 5.3 Akses berdasarkan protokol per hari.....	38
Grafik 5.4 Akses pengguna berdasarkan jenis paket	40
Grafik 5.5 Akses pengguna berdasarkan jenis produk dan provinsi	41
Grafik 5.6 Akses pengguna berdasarkan jenis produk.....	41
Grafik 5.7 Akses pengguna berdasarkan aplikasi pengirim pesan	43
Grafik 5.8 Akses pengguna berdasarkan penggunaan situs jual beli online	44
Grafik 5.9 Akses pengguna berdasarkan penggunaan situs berita online.....	45

Grafik 5.10 Akses pengguna berdasarkan penggunaan situs perjalanan online.... 46
Grafik 5.11 Akses pengguna berdasarkan penggunaan situs perbankan online.... 47
Grafik 5.12 Akses pengguna berdasarkan situs streaming film 48



DAFTAR TABEL

Tabel 2.1 Perbandingan dengan penelitian sebelumnya	9
Tabel 4.1 Penjelasan variabel dari dataset ntop.....	20
Tabel 4.2 Penjelasan variabel dari dataset pengguna	21



INTISARI

ANALISIS BIG DATA PENGGUNAAN INTERNET PADA PENGGUNA LAYANAN UBIQU

Intisari

Givan Andrian Nugroho

130707541

Seiring dengan perkembangan zaman, teknologi informasi menjadi sebuah kebutuhan utama dalam kehidupan sehari-hari terutama internet. Perkembangan internet di Indonesia berlangsung secara cepat, hal itu didukung oleh banyaknya penyedia layanan internet di Indonesia, salah satunya ada Ubiq. Ubiq menggunakan teknologi satelit VSAT-Ku Band yang dapat digunakan untuk siapa saja yang membutuhkan layanan internet khususnya di daerah-daerah terpencil. Dalam proses bisnisnya Ubiq menggunakan metode *Deep Packet Inspection* (DPI) untuk memantau lalu lintas jaringannya. Dengan menggunakan teknologi DPI, Ubiq menghasilkan data yang bervolume sangat besar atau *big data* berisikan penggunaan layanan internet oleh para penggunanya.

Dari permasalahan yang ada diatas, maka dibutuhkan sebuah proses analisis dari data-data yang dihasilkan pada proses DPI yang digabungkan dengan data pengguna yang selanjutnya data tersebut dilakukan berbagai proses-proses yang ada dalam data mining untuk mendapatkan atau menggambarkan informasi-informasi yang terdapat dalam data penggunaan internet pada layanan Ubiq tersebut. Kemudian hasil dari proses tersebut digambarkan secara visual dengan menggunakan berbagai diagram atau grafik untuk memudahkan penyampaian informasi yang didapatkan.

Dimana hasil dari pemrosesan tersebut yang berisikan informasi atau pengetahuan baru dapat digunakan oleh para pemegang kepentingan pada Ubiq untuk meningkatkan kualitas layanan dari produk Ubiq.

Kata Kunci: Ubiq, *Data Mining*, *Big Data*

Dosen Pembimbing I : Martinus Maslim, ST., MT.
Dosen Pembimbing II : Patricia Ardanari, S.Si, M.T
Dosen Penguji I : Joanna Ardhyanti Mita N, S.Kom., M.Kom
Dosen Penguji II : Stephanie Pamela Adithama, ST., MT.

Jadwal Sidang Tugas Akhir : Senin, 23 November 2020



BAB I. PENDAHULUAN

1.1. Latar Belakang

Seiring dengan perkembangan zaman, teknologi informasi merupakan sebuah kebutuhan utama di dalam kehidupan sehari-hari terutama internet. Internet, di era konvergensi digital membawa pengaruh di berbagai sektor kehidupan, menjadikan dunia tanpa batas, dan memancing munculnya berbagai teknologi pendukung, penyedia akses atau penyelenggara baru, hingga produk yang mendorong penggunaan lebih masif [1]. Internet sendiri memiliki banyak fungsi yang dapat membantu pengguna menjalankan aktivitas lebih efisien. Dengan menggunakan internet pengguna akan mendapatkan banyak hal seperti sarana penambah wawasan, komunikasi, pekerjaan, dan hiburan.

Perkembangan internet di Indonesia berlangsung secara cepat. Berdasarkan survey langsung dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) menghasilkan jumlah pengguna internet di Indonesia pada tahun 2017 sebanyak 143,26 juta jiwa atau setara dengan 54,68% dari jumlah penduduk Indonesia [2]. Berdasarkan survey tersebut itu juga diketahui bahwa penetrasi pengguna internet pada daerah timur Indonesia tergolong masih rendah di mana pada provinsi Sulawesi terdapat 48.70% jumlah penduduk yang sudah menggunakan internet disusul dengan Maluku dan Papua sebanyak 41.88%. Hal ini disebabkan karena banyaknya daerah yang belum terjangkau akses internet dikarenakan belum ada infrastruktur yang mendukung.

Ubiqu adalah layanan VSAT Ku-Band milik PT. Pasifik Satelit Nusantara yang menawarkan koneksi internet *broadband* cepat menggunakan teknologi satelit yang baru diluncurkan pada tahun 2017. Layanan Ubiqu dapat digunakan siapa saja yang membutuhkan internet,

baik di rumah, sekolah, pedesaan, rumah sakit maupun berbagai macam jenis bisnis usaha multi sektor lainnya. Ubiqum menjadi solusi terbaru layanan internet *broadband* di mana saja di seluruh wilayah Indonesia karena dapat melayani hingga ke daerah terpencil yang belum terjangkau oleh jaringan komunikasi radio, GSM, kabel dan fiber optik.

Pada proses kerjanya, Ubiqum melakukan pemrosesan data menggunakan teknologi *Deep Packet Inspection* (DPI). DPI adalah teknologi untuk memeriksa dan menganalisa keseluruhan paket dari sebuah trafik, serta mengumpulkan dan mengambil tindakan berdasarkan informasi yang dikumpulkan tersebut [3]. Selain itu DPI juga berguna untuk manajemen jaringan, keamanan jaringan, *internet censorship*, dan juga *data mining*. Dengan DPI, setiap data dari aliran paket pada trafik tersimpan ke dalam *database* sehingga menghasilkan volume data dan arus data yang sangat besar atau yang biasa disebut *big data*.

Dengan banyaknya data yang dihasilkan pada proses tersebut dibutuhkan proses analisa data. Analisis *big data* adalah proses mengumpulkan, mengatur, dan menganalisa jumlah data dalam volume besar yang bertujuan untuk menemukan informasi atau pengetahuan yang berguna. Salah satu metode untuk menganalisis big data adalah menggunakan *data mining*. *Data mining* adalah kegiatan mengekstraksi atau menambang pengetahuan dari data yang berukuran/berjumlah besar [4]. Salah satu metode dalam data mining adalah metode deskriptif analisis, Deskriptif analisis adalah metode analisis yang berfokus untuk mengetahui apa yang terjadi, dimana data yang ada diproses untuk menciptakan suatu set data historis yang dapat membantu mengungkap pola yang memberikan pengetahuan tren apa yang terjadi di masa lalu dan memberikan ide tentang apa yang saja yang mungkin terjadi di masa depan [5].

Berdasarkan penjelasan di atas, maka penelitian ini akan melakukan analisis terhadap data *deep packet inspection* yang dihasilkan dari penggunaan internet oleh para pengguna layanan Ubiqum untuk memperoleh informasi dan pengetahuan mengenai karakteristik dari para penggunanya.

Sehingga dengan diperolehnya pengetahuan tersebut dapat digunakan sebagai bahan evaluasi dan pengambilan keputusan bagi pihak Ubiq kedepannya dapat meningkatkan kualitas dari layanan Ubiq.

1.2. Rumusan Masalah

Berdasarkan penjelasan pada latar belakang diatas maka dapat dirumuskan permasalahan sebagai berikut:

1. Bagaimana cara menganalisa data pengguna ubiq berdasarkan data *deep packet inspection*?
2. Bagaimana karakteristik para pengguna layanan ubiq berdasarkan pola penggunaan internetnya?

1.3. Batasan Masalah

Batasan masalah untuk penelitian ini memiliki ruang lingkup sebagai berikut:

1. Data yang digunakan dalam penelitian ini adalah data penggunaan internet para pelanggan Ubiq yang didapatkan dari metode *deep packet inspection*.
2. Data yang digunakan dalam penelitian ini adalah data penggunaan internet pada bulan Juni tahun 2019.
3. Data pelanggan yang digunakan dalam penelitian ini hanya data provinsi tempat tinggal dan jenis produk yang digunakan.
4. Variabel yang digunakan dalam penelitian ini merupakan atribut data milik Ubiq seperti *IP address* pelanggan, *IP address* tujuan, tanggal dan *date time request*, *type request*, dan *protocol request*.
5. Data diolah menggunakan tools Jupyter Notebook dengan bahasa pemrograman Python 3

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Menganalisis pola penggunaan internet pada pelanggan Ubiq berdasarkan data *deep packet inspection*
2. Mengetahui karakteristik pada pelanggan Ubiq berdasarkan pola penggunaannya.

1.5. Metode Penelitian

Dalam penelitian ini, digunakan beberapa metode sebagai acuan dalam perencanaan, implementasi dan penelitian terhadap analisis yang akan dilakukan. Metode tersebut sebagai berikut:

1. Studi Pustaka

Studi Pustaka dilakukan dengan cara mempelajari teori-teori yang mendukung melalui buku, jurnal, artikel dan bahan pembelajaran lainnya yang berhubungan dengan analisis data, *big data*, *data mining*, dan metode lainnya yang dibutuhkan.

2. Analisis

Pada tahap analisis beberapa fase dilakukan adalah sebagai berikut:

- a. Memahami proses bisnis yang terjadi di Ubiq beserta perangkat atau *software* yang ada di dalamnya.
- b. Mengumpulkan data penggunaan internet dan data pelanggan pada layanan Ubiq
- c. Membersihkan data penggunaan internet yang akan digunakan
- d. Melakukan transformasi data untuk memudahkan proses

pemrosesan data

- e. Melakukan pemrosesan data untuk mendapatkan informasi-informasi mengenai penggunaan internet pada layanan Ubiqu
- f. Melakukan visualisasi data untuk menggambarkan secara visual informasi-informasi yang didapatkan.

3. **Penyusunan Laporan**

Pada tahap ini, penulis melakukan penyusunan laporan mengenai hasil akhir dari penelitian untuk membantu para pengambil keputusan pada Ubiqu untuk meningkatkan dan mengevaluasi kualitas layanan dari Ubiqu.

1.6. **Sistematika Penulisan**

Secara sistematis isi dari laporan ini disusun sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, metodologi penelitian, dan sistematika penulisan laporan.

BAB II TINJAUAN PUSTAKA

Bab ini berisi penjelasan dari penelitian-penelitian yang telah dilakukan sebelumnya yang berkaitan dengan penelitian yang dilakukan oleh penulis.

BAB III LANDASAN TEORI

Bab ini berisi dasar-dasar teori yang digunakan sebagai pedoman dan acuan untuk memecahkan masalah dalam penelitian.

BAB IV METODOLOGI PENELITIAN

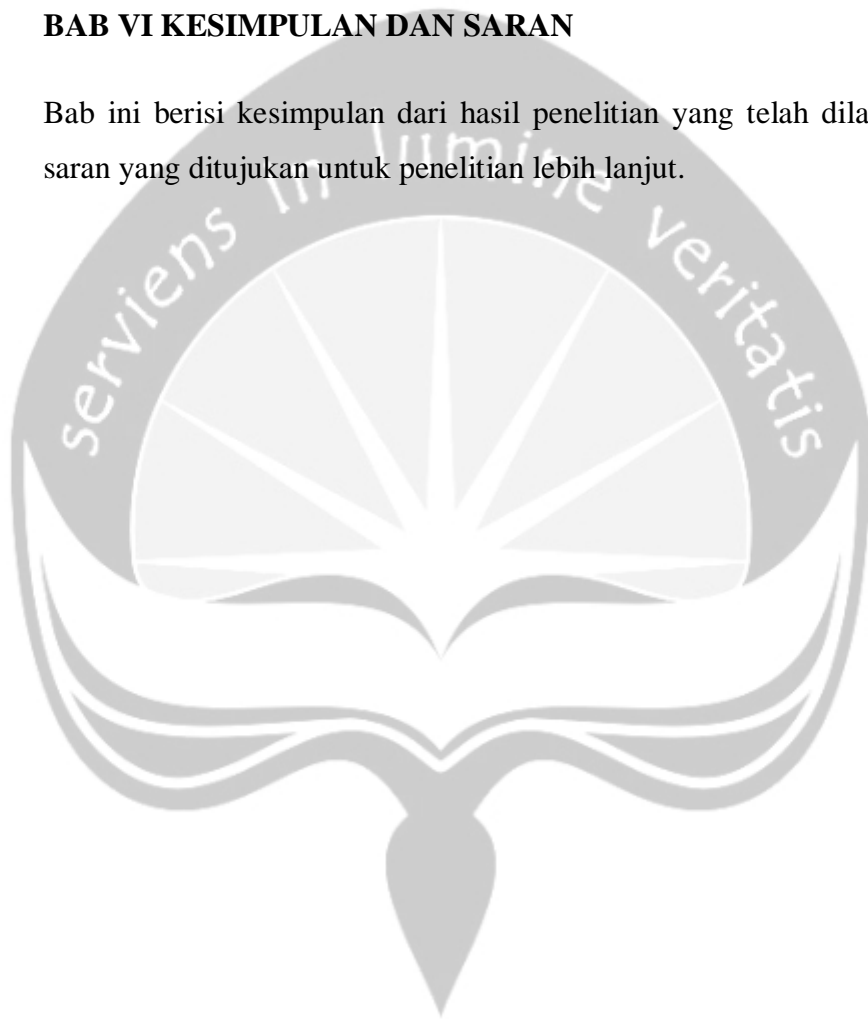
Bab ini berisi penjelasan analisis dan perancangan dari penelitian yang dilakukan.

BAB V HASIL DAN PEMBAHASAN

Bab ini berisi penjelasan mengenai hasil dan pembahasan dari penelitian yang telah dilakukan.

BAB VI KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan dan saran yang ditujukan untuk penelitian lebih lanjut.



BAB II. TINJAUAN PUSTAKA

Pada bagian ini akan dijelaskan uraian singkat mengenai pustaka yang akan digunakan oleh penulis sebagai sarana pembandingan dan acuan untuk membuat analisis, dan ulasan dari analisis-analisis yang telah dilakukan sebelumnya.

2.1 Penelitian Sebelumnya

Penelitian yang dilakukan Novri Ayu Lingga adalah penelitian terhadap Universitas Atma Jaya Yogyakarta untuk menganalisis data pengambilan mata kuliah dengan menggunakan daftar hasil studi milik alumni yang sudah lulus beberapa tahun terakhir. Sehingga data tersebut diolah dan dibersihkan terlebih dahulu yang kemudian dilakukan pengolahan data mining yang berguna untuk menemukan suatu pola atau pengetahuan yang menarik dari sekumpulan data tersebut. Dengan melakukan *clustering* K-means pada data mahasiswa program sarjana Universitas Atma Jaya Yogyakarta pada tahun akademik 2013/2013 – 2016/2017 maka penelitian tersebut menghasilkan pengelompokan dengan $k=3$ dimana *cluster* 1 merupakan kelompok yang memiliki nilai rata-rata tinggi, *cluster* 2 merupakan kelompok yang nilai rata-rata sedang dan *cluster* 3 kelompok dengan nilai rata-rata rendah. Dimana setiap *cluster*, nilai yang tertinggi dari semua kategori adalah nilai dari mata kuliah pengembangan sedangkan nilai yang terendah berasal dari mata kuliah kuantitatif [6].

Selanjutnya penelitian yang dilakukan oleh Lolavita Nakka Gasong terhadap Universitas Atma Jaya Yogyakarta untuk menganalisis pengaruh nilai seleksi jalur masuk mahasiswa yang dapat berasal dari nilai test, nilai raport, atau nilai ebtanas murni terhadap indeks prestasi kumulatif (IPK) yang didapat oleh mahasiswa tersebut. Dengan melakukan menggunakan

metode k-means untuk pengelompokan data mahasiswa berdasarkan IPK berdasarkan nilai masuk dan analisis regresi berganda untuk mengetahui apakah jalur masuk memiliki pengaruh terhadap nilai tes terhadap data mahasiswa tahun akademik 2012/2013 hingga 2016/17. Kemudian dihasilkan informasi bahwa nilai tes memiliki pengaruh sangat kecil terhadap IPK mahasiswa [7].

Penelitian selanjutnya dilakukan oleh Elvira Asril dan teman-temannya melakukan terhadap Universitas Lancang Kuning untuk menganalisis data lulusannya universitas tersebut yang terdiri dari nama mahasiswa, jurusan, sekolah asal, kota asal mahasiswa, dan indeks prestasi kumulatif untuk mendukung strategi promosi universitas tersebut. Dengan melakukan *clustering* K-means pada data lulusan tersebut menghasilkan empat *cluster*. Dimana berdasarkan *cluster* yang terbentuk strategi promosi bagi calon mahasiswa baru yang tepat sasaran untuk setiap wilayah tersebut adalah dengan mengirim tim promosi berdasarkan potensi akademik mahasiswa dengan melihat rata-rata IPK pada setiap *cluster* [8].

Dengan adanya tinjauan-tinjauan pustaka di atas, maka penulis akan melakukan analisis terhadap data penggunaan internet pada pengguna layanan Ubiqu. Dengan begitu para pihak-pihak yang bertugas mengambil keputusan pada Ubiqu dapat memperoleh informasi atau pengetahuan yang baru untuk membantu meningkatkan kualitas dan sebagai bahan evaluasi kedepannya terhadap layanan Ubiqu.

Tabel 2.1 Perbandingan dengan penelitian sebelumnya

Unsur Pembeding	[6]	[7]	[8]	(Givan, 2020)*
Judul	Analisis Data Pengambilan Mata Kuliah Program Studi Teknik Informatika Universitas Atma Jaya Yogyakarta	Analisis Pengaruh Nilai Jalur Masuk terhadap Indeks Prestasi kumulatif Mahasiswa Universitas Atma Jaya Yogyakarta	Analisis Data Lulusan dengan Data Mining untuk Mendukung Strategi Promosi Universitas Lancang Kuning	<i>Analisis Big Data Penggunaan Internet Pada Pengguna Layanan Ubiq</i>
Tujuan	Menganalisis pengambilan mata kuliah alumni Program Studi Teknik Informatika Universitas Atma Jaya Yogyakarta	Menganalisis apakah nilai jalur masuk memiliki hubungan dengan IPK mahasiswa dan mengelompokkan data mahasiswa berdasarkan IPK dan jalur masuk	Menganalisis dan mengolah data mahasiswa yang telah lulus dari Universitas Lancang Kuning untuk mendapatkan pengetahuan atau informasi penting dan bermanfaat	Menganalisis mengklasifikasi pola penggunaan internet pada pengguna layanan Ubiq
Tools	R	R	-	Python

Hasil Penelitian	Mendapatkan informasi dan pengetahuan baru mengenai karakteristik tiap cluster yang terbantu	Mendapatkan informasi bahwa nilai tes masuk berpengaruh sangat sedikit terhadap IPK	Mendapatkan informasi atau pengetahuan penting yang dapat berguna untuk strategi promosi Universitas Lancang Kuning	Mendapatkan informasi atau pengetahuan mengenai pola dan klasifikasi dari para pengguna layanan Ubiq
Sasaran Penelitian	Para pengambil keputusan di Program Studi Teknik Informatika Universitas Atma Jaya Yogyakarta	Mahasiswa Universitas Atma Jaya Yogyakarta	Para pihak pemasaran Universitas Lancang Kuning	Para pihak-pihak pengambil keputusan pada layanan Ubiq

*) penelitian yang dilakukan

BAB III. LANDASAN TEORI

3.1 Big Data

Penggunaan istilah *big data* mengacu pada evolusi dari penggunaan teknologi yang memberikan informasi yang tepat kepada pengguna pada waktu yang tepat dari kumpulan data yang tumbuh secara eksponensial sejak lama di masyarakat [5]. Selain itu big data juga sering digambarkan sebagai volume data yang besar, baik data yang tersusun secara terstruktur maupun data yang tidak terstruktur.

Fenomena big data dimulai pada tahun 2000-an di mana pada saat itu seorang analis industri yaitu Doug Laney menyampaikan karakteristik dari big data yang terdiri dari tiga bagian penting, yang diantaranya:

a. Volume

Mewakili jumlah data yang dihasilkan, disimpan, dan dioperasikan dalam sistem. Peningkatan pada volume dijelaskan oleh peningkatan pada jumlah data yang dihasilkan dan disimpan, tetapi juga oleh kebutuhan untuk mengeksploitasinya.

b. Variasi

Mewakili data yang dikumpulkan mempunyai banyak format yang berbeda-beda. Dimulai dari data terstruktur, data numerik, data dokumen, data email, video, audio, dan lain-lainnya.

c. Kecepatan

Mewakili aliran data data yang harus ditangani dengan secara cepat dan tepat yang dapat dilakukan melalui hardware atau software.

3.2 Data Mining

Data mining adalah salah satu langkah penting dalam proses *knowledge discovery in database* (KDD) yang menghasilkan pola atau model dari data. Data mining digunakan untuk menemukan pola baru dari banyak data dalam database dan berfokus pada algoritma untuk mendapatkan pengetahuan yang bermanfaat [9].

Terdapat beberapa fungsi dalam data mining, salah satunya adalah fungsi deskriptif. Fungsi ini berguna untuk mendeskripsikan atau menggambarkan informasi berbasis pengetahuan yang bermanfaat dan yang belum diketahui dari sekumpulan data. Pada penelitian kali ini fungsi tersebut yang akan digunakan untuk mengetahui informasi-informasi yang belum diketahui dari data penggunaan internet pada layanan Ubiqu.

Data mining juga memiliki tujuan untuk mengekstraksi serta mengidentifikasi suatu data demi informasi tertentu yang berhubungan dengan suatu database besar atau big data. Selain itu data mining juga memiliki beberapa fungsi dasar sebagai berikut:

a. *Classification*

Klasifikasi pada data mining berguna untuk menemukan model yang dapat menganalisis dan mengklasifikasikan sebuah data menjadi beberapa kelas yang sudah ditentukan.

b. *Regression*

Regresi pada data mining berguna untuk memetakan sebuah data ke variabel

prediksi bernilai nyata.

c. *Clustering*

Klasterisasi pada data mining berguna untuk mengidentifikasi seperangkat kategori atau kluster terbatas untuk menggambarkan data.

d. *Dependency Modeling*

Dependency Modeling berguna untuk menemukan model yang menggambarkan ketergantungan yang signifikan antara variabel.

e. *Deviation Detection*

Deviation Detection berguna untuk menemukan perubahan yang paling signifikan pada data.

f. *Summarization*

Summarization berguna untuk menemukan deskripsi singkat dari subset data.

3.3 Reverse DNS Lookup

Reverse DNS Lookup (rDNS) adalah salah satu teknologi yang terdapat di *Domain Name System* (DNS) yang berguna untuk menemukan nama domain dari suatu IP address. Proses translasi tersebut dilakukan menggunakan PTR records, dimana pada umumnya proses rDNS dapat digunakan untuk melacak nama domain pada sebuah IP address untuk mengidentifikasi pengirim surel spam, atau domain name dari komputer yang mencoba melakukan peretasan pada sistem.

3.4 Pemrograman Python

Python adalah bahasa pemrograman yang pertama kali muncul pada tahun 1991 yang dirancang oleh seorang bernama Guido Van Rossum. *Python* dapat disebut sebagai interpretatif multiguna, tidak seperti bahasa pemrograman lainnya yang sulit untuk dibaca dan dipahami, *python* lebih mementingkan pembacaan kode agar lebih mudah untuk memahami sintaks yang ada sehingga dapat dengan mudah untuk digunakan ulang.

Pada tahun 2018 menjadi bahasa pemrograman yang paling populer dikarenakan kemudahan dan kepraktisan *python* dan *python* juga menjadi bahasa paling populer dalam pengelolaan data. *Python* banyak digunakan dalam pengelolaan data dikarenakan banyaknya library pada *python* yang mendukung untuk melakukan pengelolaan, pemrosesan, visualisasi, sampai pengimplementasian machine learning pada data.

3.5 Ubiq

Ubiq adalah layanan internet satelit VSAT Broadband yang dikembangkan oleh PT Pasifik Satelit Nusantara dimana ubiq menyediakan layanan internet menggunakan pemancar antena yang tersedia untuk seluruh Indonesia terutama untuk wilayah tertinggal. Ubiq hadir sebagai solusi koneksi internet untuk di rumah, sekolahan, pedesaan, klinik, puskesmas dan lainnya yang tidak ada jaringan komunikasi terestrial (kabel, ADSL, fiber optik, radio, atau GSM).

Dalam proses bisnisnya, Ubiq menyediakan berapa paket yang dibagi menjadi paket FIT dan BIG. Paket FIT adalah jenis paket yang berbasis kuota terdiri dari FIT 1 yang berisi kuota 8-14 GB, FIT 2 yang berisi kuota 14 - 28, FIT 2 yang berisi kuota 24 - 50 GB, dan FIT 4 yang berisi 48 - 100 GB per bulannya tergantung dari wilayah tempat

pemasangannya. Sedangkan untuk paket BIG adalah jenis paket yang unlimited dengan Fair Usage Policy (FUP) atau kuota dengan batas penggunaan wajar, paket BIG sendiri terdiri dari BIG 1 yang berisi kuota 104 - 220 GB, BIG 2 yang berisi kuota 166 - 370 GB, dan BIG 3 yang berisi kuota 320 - 680 GB per bulannya tergantung dari wilayah tempat pemasangannya.

Selain kedua paket tersebut, Ubiq memiliki layanan lainnya yang bernama Ubiq Sinyalku, Ubiq Sinyalku adalah layanan akses internet melalui konektivitas WiFi menggunakan akses satelit broadband yang dapat dipasang di mana saja di seluruh Indonesia. Ubiq Sinyalku memungkinkan penggunanya untuk membuka bisnis WiFi internet dan juga dapat mengembangkan bisnis para penggunaannya dengan cara menyewakan akses konektivitas internet kepada orang lain.

3.6. Deep Packet Inspection

Deep Packet Inspection (DPI) adalah sebuah teknologi yang memungkinkan sebuah aplikasi atau protokol yang berfungsi untuk memfilter jaringan. DPI melakukan evaluasi pada bagian data dan header dalam setiap paket yang ada di jaringan yang digunakan untuk memfilter protokol, spam, virus, atau berbagai gangguan lainnya yang ditentukan pada setiap jaringan yang ada.

Pada proses kerjanya, DPI melakukan filtering yang biasanya berada di dalam sebuah jaringan dimana pada proses tersebut DPI mengevaluasi konten dari setiap paket yang melewati *firewall* sesuai dengan peraturan yang ditentukan oleh *Internet Service Provider*, atau *network administrator*. Selain itu DPI juga menentukan apa yang dilakukan pada paket dalam sebuah jaringan.

Selain itu DPI juga bekerja dengan memeriksa paket hingga application layer pada OSI Layer untuk mendapatkan informasi jenis trafik

[10]. Hal itu digunakan untuk mengatasi berbagai masalah identifikasi paket data yang berdasarkan hanya dengan port number dari paket tersebut.

3.7. Ntop

Ntop adalah program komputer yang berfungsi untuk memantau lalu lintas penggunaan jaringan menggunakan metode *deep packet inspection*. Aplikasi Ntop dapat menampilkan informasi traffic dan daftar pengguna yang menggunakan sebuah jaringan tertentu secara mendetail. Selain itu Ntop juga memiliki beberapa kegunaan lainnya seperti:

1. Untuk mengklasifikan lalu lintas pada jaringan berdasarkan protokol yang digunakan
2. Menampilkan statistik dari lalu lintas pada jaringan
3. Menyimpan lalu lintas jaringan ke dalam *database*.
4. Mengidentifikasi pengguna atau host yang menggunakan jaringan
5. Mengidentifikasi setiap *flow* atau *request* yang terjadi pada jaringan

BAB VI. PENUTUP

6.1. Kesimpulan

Berdasarkan dari pembahasan dari bab-bab sebelumnya maka didapat beberapa kesimpulan sebagai berikut:

1. Proses analisis pola penggunaan internet pada pelanggan Ubiq berdasarkan data *deep packet inspection* berhasil dilakukan dengan dilakukan beberapa fase seperti pengumpulan, pembersihan, transformasi, pemrosesan, hingga visualisasi data dengan menggunakan bahasa pemrograman Python.
2. Karakteristik dari para pengguna layanan Ubiq yang umumnya berada di daerah rural tidak jauh berbeda dengan masyarakat di daerah urban dilihat dari pola penggunaan jaringan *internetnya*. Selain itu pengguna layanan Ubiq yang aktif pada bulan Juni 2019 paling banyak berasal dari kawasan Indonesia bagian timur khususnya dari Provinsi Papua dan pada layanan Ubiq, produk Sinyalku lebih banyak pelanggan aktifnya dibandingkan dengan Ubiq.

6.2. Saran

Berdasarkan dari pembahasan dari bab-bab sebelumnya maka didapat beberapa saran sebagai berikut:

1. Informasi-informasi yang sudah didapatkan dari proses analisis dalam penelitian dapat digunakan sebaik-baiknya untuk mengambil keputusan di masa depan oleh para pemegang kepentingan pada layanan Ubiq
2. Proses analisa penggunaan internet menggunakan metode *deep packet inspection* bisa lebih mendetail dengan menggunakan

waktu pada saat sebuah akses dilakukan sehingga bisa mengetahui pola penggunaan pada pagi, siang, atau malam hari



DAFTAR PUSTAKA

- [1] J. Curran, N. Fenton and D. Freedman, "Misunderstanding The Internet," *Pedagoges: AN International Jurnal*, no. 11, pp. 270-277, 20316.
- [2] Asosiasi Penyelenggara Jasa Internet Indonesia, "Hasil Survei Penetrasi dan Perilaku Pengguna Internet Indonesia," 2017.
- [3] F. E. Nugroho, G. B. Satrya and T. Brotoharsono, "Analisis Perbandingan Performasi Deep Packet Inspection Firewall Antara L7-Filter dan nDPI," *e-Procceeding of Engineering*, vol. 2, p. 1469, 2015.
- [4] H. Susanto and S. , "Data Mining Untuk Memprediksi Prestasi Siswa Berdasarkan Sosial Ekonomi, Motivasi, Kedisiplinan Dan Prestasi Masa Lalu," *Jurnal Pendidikan Vokasi*, vol. 5, no. 2, p. 225, 2014.
- [5] Y. Riahi and S. Riahi, "Big Data and Big Data Analytics: Concepts, Types and Technologies," *International Journal of Research and Engineering*, vol. 5, no. 9, p. 525, 2018.
- [6] N. A. Lingga, "Analisis Pengambilan Mata Kuliah Program Studi Teknik Informatika Universitas Atma Jaya Yogyakarta," *Tugas Akhir*, 2019.
- [7] L. N. Gasong, "Analisis Pengaruh Nilai Jalur Masuk Terhadap Indeks Prestasi Kumulatif Mahasiswa Unviersitas Atma Jaya Yogyakarta," *Tugas Akhir*, 2019.
- [8] E. Asril, F. Wiza and Y. Yunefri, "Analisis Data Lulusan dengan Data Mining untuk Mendukung Strategi Promosi Universitas Lancang Kuning," *Tugas Akhir*, 2015.
- [9] T. Silawattananusarn and K. Tuamsuk, "Data Mining and Its Applications for Knowledge Management: A Literature Review from 2007 to 2012," *International Journal of Data Mining & Knowledge Management Process (IJDKP)*, vol. 2, no. 5, pp. 13-24, 2012.

- [10] T. S. A. Valianta and D. Stiawan, "Klasifikasi Trafik Terenkripsi Menggunakan Metode Deep Packet Inspection (DPI)," *ANNUAL RESEARCH SEMINAR 2016*, vol. 2, no. 1, p. 424, 2016.

