

# **ANALISIS POPULARITAS OBJEK WISATA KABUPATEN/KOTA DI BALI MENGGUNAKAN TEKNOLOGI BIG DATA**

## **TUGAS AKHIR**

Diajukan guna memenuhi sebagian persyaratan dalam rangka menyelesaikan  
Pendidikan Sarjana Strata Satu (S1) Program Studi Teknologi Informasi



**NI PUTU AYU WIDIARI  
NIM: 1605551068**

**PROGRAM STUDI TEKNOLOGI INFORMASI  
FAKULTAS TEKNIK  
UNIVERSITAS UDAYANA  
2019**

# **BAB I**

## **PENDAHULUAN**

Bab I berisi pembahasan mengenai pendahuluan dari penelitian. Isi bab I berupa latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, dan sistematika penulisan.

### **1.1 Latar Belakang**

Sejak didirikan pada 2006, Twitter dengan cepat menjadi salah satu layanan paling populer di lanskap media sosial. Informasi yang beredar juga sangat beragam mulai dari berita, opini, kritik yang bersifat positif, negatif, ataupun netral. Perusahaan twitter menyatakan memiliki 302 juta pengguna aktif bulanan, yang menulis sekitar 500 juta posting per hari pada Juni 2015. Pengguna Twitter 80% menggunakan layanan dari perangkat seluler. Twitter mendukung lebih dari 35 bahasa. Akun terdaftar 77% di luar AS dimana pos Twitter tersedia dalam berbagai bahasa (sumber: <https://about.twitter.com/company>, 1 Juni 2015). Twitter sebagai sebuah situs jejaring sosial memberikan akses kepada penggunanya untuk mengirimkan sebuah pesan singkat yang terdiri dari maksimal 140 karakter (disebut *tweet*).

Pulau Bali dikenal sebagai pulau yang penuh dengan objek wisata yang mampu menarik minat para wisatawan mancanegara atau domestik. Setiap objek wisata memiliki potensinya masing-masing yang membedakan satu wisata dengan wisata lainnya yang biasanya diketahui dengan dilakukan survei secara langsung, namun, untuk melakukan hal tersebut membutuhkan waktu yang lama dan tidak praktis. Keberagaman data objek wisata menyebabkan jumlah dan variasi data menjadi berlimpah yang dikenal dengan istilah *big data*. Data yang berlimpah dan beragam tersebut tentunya memerlukan suatu proses pengolahan di dalamnya menggunakan Teknologi Big Data.

Berkaitan dengan permasalahan yang dikemukakan di atas, melalui penelitian ini akan diupayakan suatu analisis sosial media yaitu twitter untuk mengetahui analisis

tren obyek wisata di tiap kabupaten/kota pulau Bali. Kabupaten/kota yang akan dianalisis dikerucutkan menjadi 6 lokasi, yaitu Tabanan, Karangasem, Gianyar, Denpasar, Badung, dan Nusa Penida. Data twitter yang dikumpulkan menggunakan Twitter API yang berisikan hastag(#) tiap objek wisata. Objek wisata pada tiap kabupaten/kota tersebut adalah Tabanan (#tanahlot, #batukaru, #Bedugul, #beratanlake), Karangasem (#besakih, #candidasa, #lempuyang), Gianyar (#monkeyforest, #sukawati, #ubud), Denpasar (#sanur, #kertalangu), Badung (#GWK, #Kuta, #puratamanayun, #uluwatu, #sangeh), Nusa Penida (#brokenbeach, #crystalbay, #kelingkingbeach). Data yang dikumpulkan akan mengalami proses penyimpanan, pengolahan, dan analisis data sebelum data divualisasi dalam bentuk grafik/peta/diagram. Platform/tools yang digunakan dalam analisis data twitter adalah Apache Spark dan program Python.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang tersebut, beberapa rumusan masalah yang dapat diambil antara lain sebagai berikut.

1. Bagaimana cara mengolah serta menganalisis data twitter objek wisata dengan program python.
2. Bagaimana pengaruh media sosial twitter terhadap pengaruh minat wisatawan terhadap objek wisata di Bali.

## **1.3 Tujuan**

Berdasarkan rumusan masalah yang telah dipaparkan sebelumnya, tujuan dari penulisan laporan Tugas Akhir antara lain sebagai berikut.

1. Dapat mengembangkan *prototype* sistem yang dapat memudahkan pengumpulan, pengolahan, serta analisis tweet dalam jumlah besar.
2. Dapat mengetahui tren dan popularitas objek wisata Pulau Bali dari hasil visualisasi *prototype* sistem.

#### **1.4 Manfaat**

Berdasarkan rumusan masalah yang telah dipaparkan sebelumnya, manfaat dari penulisan laporan Tugas Akhir antara lain sebagai berikut.

1. Hasil penelitian tugas akhir ini diharapkan dapat menambah bahan informasi bagi masyarakat luas yang berkaitan dengan analisis big data pada *twitter* terkait potensi objek wisata di pulau Bali.
2. Hasil penelitian tugas akhir ini diharapkan dapat mengembangkan prototype sistem yang dapat memudahkan pengumpulan, pengolahan, serta analisis tweet dalam jumlah besar yang dapat digunakan untuk penelitian akademik ataupun analisis trend.

#### **1.5 Batasan Masalah**

Batasan masalah dalam penelitian ini bertujuan untuk menghindari ruang lingkup pembahasan yang terlalu luas dan memudahkan dalam penyelesaian masalah sesuai tujuan yang ingin dicapai. Batasan masalah dalam penelitian ini adalah sebagai berikut:

- 1) Penelitian ini hanya berfokus pada pembuatan analisis popularitas objek wisata pada kabupaten/kota di Bali dengan Teknologi Big Data.
- 2) Analisis yang dilakukan menggunakan tools pengolahan yaitu MongoDB sebagai tools penyimpanan, Apache Spark sebagai tools analisis data, dan Tableau sebagai tools visualisasi.

#### **1.6 Sistematika Penulisan**

Sistematika penulisan yang digunakan pada laporan Tugas Akhir ini terdiri dari pembahasan umum secara teori kemudian dilengkapi dengan pemaparan contoh dari setiap penjelasan tersebut.

#### **BAB I : Pendahuluan**

Pada bagian ini dijelaskan mengenai ide dasar Analisis Popularitas Objek Wisata Kabupaten/Kota Di Bali Menggunakan Teknologi Big

Data yang meliputi latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah serta sistematika penulisan.

**BAB II : Tinjauan Pustaka**

Pada bagian ini membahas tentang landasan teori secara umum yang menjadikan dasar untuk digunakan dalam proses Analisis Popularitas Objek Wisata Kabupaten/Kota Di Bali Menggunakan Teknologi Big Data

**BAB III : Metodologi Penelitian**

Pada bagian ini dijelaskan mengenai metode yang digunakan dalam penelitian serta proses pembuatan laporan yang meliputi tempat dan waktu penelitian, alur penelitian, pemodelan sistem, perancangan sistem dan jadwal kegiatan.

**BAB IV : Pembahasan**

Pada bagian ini dijelaskan mengenai aplikasi berupa rancangan yang telah dibuat, pengujian aplikasi serta hasil dari pengujian.

**BAB V : Penutup**

Pada bagian ini dijelaskan mengenai kesimpulan yang mengacu pada rumusan masalah dan tujuan penelitian serta saran untuk pengembangan aplikasi selanjutnya.

## **BAB II**

### **TINJAUAN PUSTAKA**

Bab ini akan membahas tentang penelitian sebelumnya (*State of the Art*) dan dasar teori yang akan dijadikan acuan atau landasan dalam pengerjaan Tugas Akhir ini.

#### **2.1     *State of The Art***

Penerapan Teknologi Big Data dalam analisis data twitter pernah dilakukan sebelumnya, sehingga menjadi acuan untuk membuat suatu sistem yang dapat melakukan aktivitas lebih dari sistem sebelumnya.

Salah satu penelitian yang telah dilakukan tertera pada jurnal berjudul “Sentiment Analysis on Twitter Data using Apache Spark Framework”. Cara kerja sistem tersebut adalah dengan mengumpulkan data dari social media berupa twitter. Data twitter yang telah dikumpulkan akan digunakan untuk sentiment analisis untuk beberapa permasalahan/opini masyarakat. Data twitter tersebut kemudian akan mengalami preprocessing dari beberapa fitur Machine Learning agar hasil yang diperoleh lebih akurat pada proses klasifikasi Sentimen Analisis. Jurnal ini mengangkat sentiment analisis dengan bantuan tools berupa Apache Spark dengan metode klasifikasi sentiment yang berbeda yaitu menggunakan algoritma naïve bayes, logitdic regression, dan decision tree. Data yang diperoleh akan divisualisasikan dalam bentuk diagram batang untuk menampilkan perbedaan akurasi dari ketiga metode klasifikasi sentiment tersebut (Hossam Elzayady,dkk,2018).

Contoh lainnya adalah pada jurnal dengan judul “Analisis Sentimen Terhadap Pemerintahan Joko Widodo Pada Media Sosial Twitter Menggunakan Algoritma Naives Bayes Classifier”. Pada jurnal tersebut dijelaskan bahwa untuk menggali informasi dan melakukan klasifikasi teks diperlukan analisis sentimen yaitu sentiment positif dan negatif. Topik yang diangkat adalah opini masyarakat mengenai pemerintahan Joko Widodo. Data opini diperoleh dari jejaring sosial Twitter berupa

tweet berdasarkan query dalam Bahasa Indonesia. Data yang digunakan berjumlah 400 tweet terdiri dari 300 data latih dan 100 data uji. Data latih merupakan data yang telah diketahui sentimennya, 300 data latih terdiri dari 150 data kelas sentimen negatif dan 150 data kelas sentimen positif. Penelitian ini bertujuan untuk menentukan tweet merupakan tweet positif atau negatif yang disampaikan di Twitter dalam Bahasa Indonesia (Yonathan Sari Mahardhika dan Eri Zuliarso, 2018).

Lalu contoh lainnya ada pada jurnal dengan judul “Twitter Streaming API Using Apache Spark in Big Data Analytics”. Pada jurnal ini, dijelaskan bahwa untuk menganalisis jumlah tweet yang lebih besar dari twitter untuk mendapatkan pola yang berbeda dan mengekstrak informasi yang relevan adalah tantangan besar. Apache spark adalah platform yang digunakan untuk menganalisis data besar secara efisien. Untuk mengetahui konteks yang digunakan, dibuatkan pola yang relevan dari data besar dengan menggunakan Twitter Streaming API. Jurnal ini mengeksplorasi konsep API streaming twitter untuk mengekstrak informasi yang relevan dari sejumlah besar data dari tweet twitter (Sakshi dan Shuchita Upadhyaya, 2016).

Kemudian, contoh lainnya ada pada jurnal dengan judul “Action Rules for Sentiment Analysis on Twitter Data using Spark”. Sistem ini merupakan Aturan Tindakan Khusus berdasarkan algoritma Grabbing strategy (SARGS). Untuk ini, dilakukan analisis komparatif metaaksi yang menghasilkan implementasi algoritmik di sistem yang didorong Apache Spark dan sistem yang digerakkan Hadoop konvensional menggunakan data jejaring sosial Twitter dan mengevaluasi hasilnya. Analisis Sentimental dilakukan berdasarkan data jaringan sosial berbasis corpus, dan menguji total waktu yang dibutuhkan oleh sistem dan sub komponennya untuk pemrosesan data. Hasil menunjukkan waktu komputasi yang lebih cepat untuk sistem Spark dibandingkan dengan Hadoop MapReduce untuk menerapkan metode pembuatan meta-action (Jaishree Ranganathan, dkk, 2017).

## **2.2 Twitter**

Twitter, salah satu layanan microblogging terpopuler, telah lebih dari 200 juta pengguna dan menghasilkan 110 juta tweets (mis., 140- pesan teks karakter) setiap hari (Chiang, 2011). Itu memungkinkan pengguna untuk menulis pesan pendek dan menyiarkannya ke *follower* mereka, dikenal sebagai pengikut, dalam waktu nyata. Dibandingkan dengan blogging tradisional, microblogging menekankan kecepatan dan singkatnya, fokus pada hal-hal terjadi sekarang (Oulasvirta,dkk,2010). Beberapa penelitian menunjukkan bahwa sebagian besar pengguna menggunakan microblog untuk menggambarkan rutinitas harian mereka, melakukan percakapan, melaporkan berita, dan berbagi informasi (Java,dkk, 2007; Naaman, Boase, & Lai, 2010), memproduksi berbasis teks ‘Aliran kesadaran sosial’ yang terdiri dari pemikiran sehari-hari, perasaan, dan percakapan (Naaman,dkk, 2010).

## **2.3 Obyek Wisata**

Menurut UU RI No 9 Tahun 1990 tentang Kepariwisata, dinyatakan bahwa obyek dan daya tarik wisata adalah segala sesuatu yang menjadi sasaran wisata baik itu pembangunan obyek dan daya tarik wisata, yang dilakukan dengan cara mengusahakan, mengelola dan membuat obyek-obyek baru sebagai obyek dan daya tarik wisata. Dalam undang-undang di atas, yang termasuk obyek dan daya tarik wisata terdiri dari:

1. Obyek dan daya tarik wisata ciptaan Tuhan Yang Maha Esa, yang berwujud keadaan alam serta flora dan fauna, seperti: pemandangan alam, panorama indah, hutan rimba dengan tumbuhan hutan tropis serta binatang-binatang langka.
2. Obyek dan daya tarik wisata hasil karya manusia yang berwujud museum, peninggalan purbakala, peninggalan sejarah, seni budaya, pertanian (wisata agro), wisata tirta (air), wisata petualangan, taman rekreasi, dan tempat hiburan lainnya.



3. Sasaran wisata minat khusus, seperti: berburu, mendaki gunung, gua, industri dan kerajinan, tempat perbelanjaan, sungai air deras, tempat-tempat ibadah, tempat- tempat ziarah, dan lain-lain.
4. Pariwisata adalah segala sesuatu yang berhubungan dengan wisata, termasuk pengusahaan objek dan daya tarik wisata serta usaha-usaha yang terkait di bidang tersebut. Dengan demikian pariwisata meliputi Semua kegiatan yang berhubungan dengan perjalanan wisata.

## **2.4 Bahasa Pemrograman Python**

*Python* adalah bahasa pemrograman interpretatif multiguna dengan filosofi perancangan yang berfokus pada tingkat keterbacaan kode *Python* diklaim sebagai bahasa yang menggabungkan kapabilitas, kemampuan, dengan sintaksis kode yang sangat jelas, dan dilengkapi dengan fungsionalitas pustaka standar yang besar serta komprehensif. *Python* juga didukung oleh komunitas yang besar .Bahasa pemrograman *Python* digunakan karena bahasa pemrograman ini hampir mendekati bahasa manusia. *Python* juga akan digunakan dalam pembuatan GUI agar sistem yang dibuat lebih mudah dioperasikan (Abed, A.A. and S.A. Rahman).

## **2.5 Apache Spark**

Apache Spark adalah salah satu platform pemrosesan sumber terbuka terdistribusi open-source yang lebih baru, kilat dan cepat berdasarkan prinsip yang sama dengan Hadoop, yang dikembangkan dengan cara untuk mengembangkan kecepatan komputasi. Hadoop memiliki beberapa masalah yang berkaitan dengan kinerja dalam kasus tertentu, seperti tugas algoritme berbasis grafik atau tugas berulang. Namun, Hadoop tidak bisa menyimpan data menengah untuk mendapatkan kinerja tinggi; itu mengalir data ke disk antara setiap langkah membaca dan menulis. Spark meminimalkan jumlah siklus baca dan tulis sehingga 10 kali lebih cepat daripada Hadoop pada disk dalam aplikasi yang berjalan; juga, menyimpan data antara dalam

memori membuat Spark 100 kali lebih cepat ketika berhadapan dengan memori (sumber: <http://spark.apache.org>, 11 April 2019)

Data terdistribusi tangguh set (RDD) mungkin merupakan ide teoretis utama dalam percikan, yang merepresentasikan kumpulan objek hanya-baca yang terbagi melintasi sekumpulan node, dengan kemampuan untuk membangunnya kembali jika muncul kejadian kehilangan partisi apa pun (Hamstra,dkk,2013). Spark terkenal karena kapasitasnya untuk menyelesaikan batch, interaktif, dan pembelajaran mesin selain streaming semua dalam kelompok yang sama. Salah satu fitur utama Spark adalah skalabilitasnya, sehingga cluster dapat dilengkapi dengan sejumlah n node. Fitur utama lain dalam Spark adalah fleksibilitas bahasa sehingga pengembang Spark dapat menggunakan (API) dalam bahasa pemrograman Scala, Python, Java dan R.

## **2.6 MongoDB**

MongoDB adalah *database* tujuan yang umum yang kuat, fleksibel, dan dapat diukur. MongoDB menggabungkan kemampuan untuk menskala dengan fitur seperti indeks sekunder, kueri rentang, penyortiran, agregasi, dan indeks geospasial. MongoDB adalah basis data berorientasi dokumen, bukan basis data relasional. Alasan utama untuk menjauh dari model relasional adalah untuk membuat penskalaan menjadi lebih mudah, tetapi ada beberapa keuntungan lain juga. Sebuah database berorientasi dokumen menggantikan konsep "baris" dengan yang lebih fleksibel model, "dokumen." Dengan memungkinkan dokumen dan array tersemat, pendekatan berorientasi dokumen memungkinkan untuk merepresentasikan hubungan hierarki yang kompleks dengan satu catatan. Ini secara alami cocok dengan cara pengembang dalam berorientasi objek modern bahasa memikirkan data mereka (Kristina Chodorow, 2013)

## **2.7 Tableau**

Tableau adalah *software* yang membantu manajemen dalam mempercepat pengambilan keputusan berdasarkan visualisasi data. Data Visualisasi dalam penelitian ini akan ditampilkan dalam bentuk diagram batang dan map atau peta.

## **BAB III**

### **METODOLOGI DAN PERANCANGAN SISTEM**

Bab ini berisikan mengenai gambaran umum dari sistem yang akan dibuat, serta metode yang digunakan dalam penyelesaian penelitian Tugas Akhir.

#### **3.1 Waktu Penelitian**

Pembuatan tugas akhir ini dilakukan mulai pada bulan Maret 2019 hingga bulan Maret 2020. Penelitian ini akan berjalan selama kurang lebih setahun agar data yang dikumpulkan lebih banyak dan maksimal dalam proses pengerjaan.

#### **3.2 Sumber Data**

Data yang digunakan dalam penelitian ini adalah data objek wisata yang diperoleh data sosial media Twitter yang diambil secara berkala dari bulan Maret 2019 dengan mengakses Twitter API. Data twitter yang dikumpulkan berisikan hastag(#) tiap objek wisata kabupaten/kota di Bali. Objek wisata pada tiap kabupaten/kota tersebut adalah Tabanan (#tanahlot, #batukaru, #Bedugul, #beratanlake), Karangasem (#besakih, #candidasa, #lempuyang), Gianyar (#monkeyforest, #sukawati, #ubud), Denpasar (#sanur, #kertalangu), Badung (#GWK, #Kuta, #puratamanayun, #uluwatu, #sangeh), Nusa Penida (#brokenbeach, #crystalbay, #kelingkingbeach).

#### **3.3 Metode Pengumpulan Data**

Penelitian ini menggunakan dua buah metode dalam pengumpulan data, agar mendapatkan data yang benar-benar akurat dan relevan, yaitu sebagai berikut.

##### **3.3.1 Studi Pustaka**

Studi pustaka adalah pengumpulan data yang penulis ambil dari berbagai macam buku-buku, literatur, referensi dan dari berbagai data yang bersumber dari media global, seperti internet yang berkaitan dengan penulisan tugas akhir ini dan juga

dapat mendukung penelitian yang penulis buat. Selanjutnya, dengan cara mempelajari dan memahami sistem yang berhubungan dengan masalah yang akan dibahas dalam penelitian ini.

### **3.4 Instrumen Perancangan dan Pembuatan Sistem**

Tahapan ini akan menganalisa beberapa kebutuhan yang diperlukan dalam pengerjaan analisis popularitas objek wisata kabupaten/kota di Bali menggunakan Teknologi Big Data, tahapan ini terdiri dari analisis kebutuhan perangkat lunak (*software*) dan kebutuhan perangkat keras (*hardware*).

#### **3.4.1 Kebutuhan Perangkat Lunak**

Menganalisa kebutuhan perangkat lunak yang diperlukan dalam merancang analisis popularitas objek wisata kabupaten/kota di Bali menggunakan Teknologi Big Data, antara lain *operating system* yang digunakan adalah Windows 10 dan JetBrains PyCharm Community Edition 2018.2.2 x64.

#### **3.4.2 Kebutuhan Perangkat Keras**

Tahapan ini merupakan tahap perancangan dan desain dari perangkat-perangkat yang ada dalam pembuatan analisis popularitas objek wisata kabupaten/kota di Bali menggunakan Teknologi Big Data, yaitu sebagai berikut.

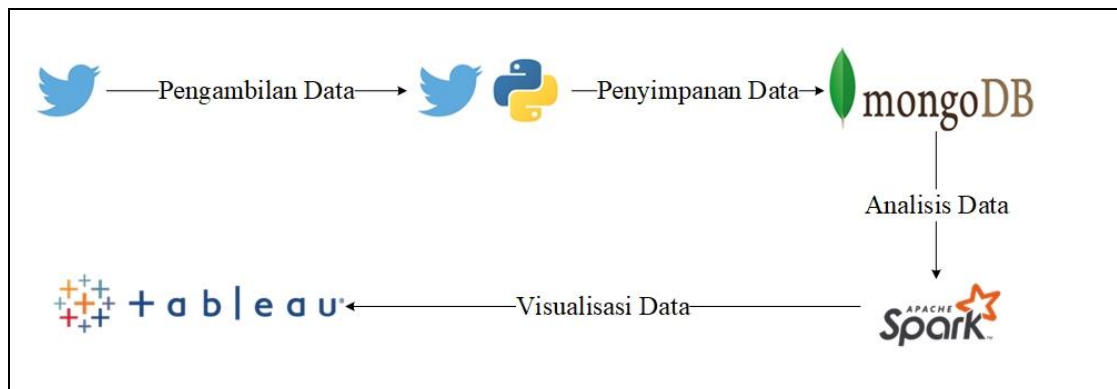
1. Laptop dengan spesifikasi Intel Core i5-4210U
2. RAM 4GB
3. Hard Disk 1TB, dan
4. VGA Nvidia GeForce GT 930X

### **3.5 Perancangan Sistem**

Tahapan ini merupakan tahap perancangan dan desain dari perangkat-perangkat yang ada dalam pembuatan analisis popularitas objek wisata kabupaten/kota di Bali menggunakan Teknologi Big Data. Tahap ini terdiri atas gambaran umum sistem dan desain algoritma, yang akan dipaparkan berikut ini.

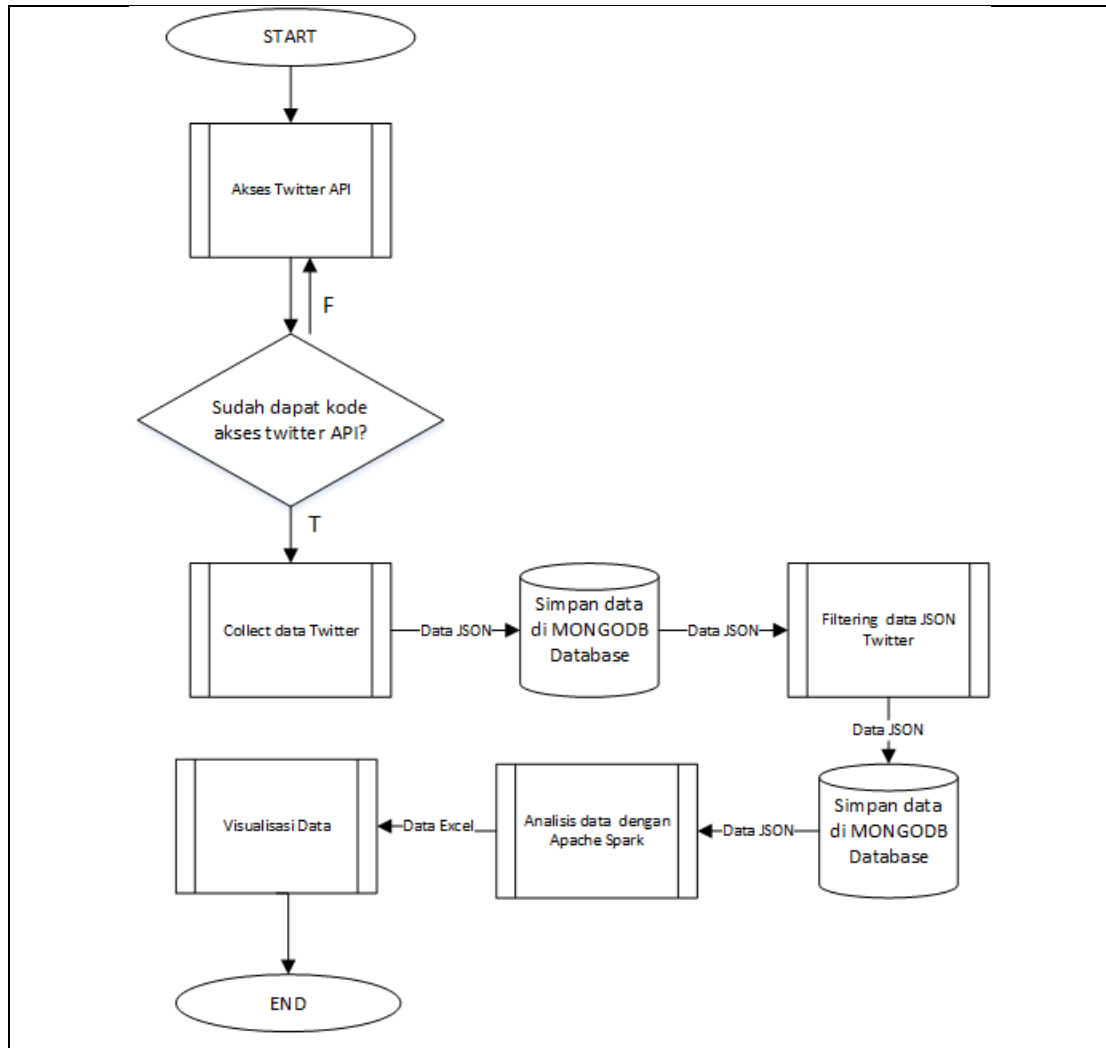
### 3.5.1 Gambaran Umum Sistem

Project Akhir Teknologi Big Data yang akan dikembangkan ini memungkinkan. Gambaran umum sistem dari Project Akhir Teknologi Big Data dapat dilihat pada Gambar 3.1



**Gambar 3.1** Gambaran Umum Sistem

Gambar 3.1 merupakan gambaran umum sistem dimana, pengambilan data melalui social media twitter, data yang telah diambil akan disimpan dalam mongodb datastore, data yang disimpan akan mengalami proses cleaning dan analisis menggunakan tools apache spark. Data bersih yang telah mengalami proses cleaning dan analisis akan divisualisasi dalam bentuk grafik menggunakan tableau.



**Gambar 3.2** Flowchart Analisis Objek Wisata

Gambar 3.2 merupakan flowchart sistem Analisis Objek Wisata yang akan dibuat yang terdiri dari empat tahapan utama yaitu Collect Data Twitter, Filtering Data, Analisis Data Twitter menggunakan Apache Spark, dan Visualisasi Data. Sebelum tahap collect data, diharuskan untuk mengakses twitter API dengan mendaftar pada <https://developer.twitter.com/en/apps> dan mendapat kode API yang akan digunakan untuk collect data twitter.

### 3.5.1 Collect Data Twitter

Collect data twitter merupakan tahapan pertama dari proses analisis data twitter objek wisata. Tahapan ini menggunakan kode token yang telah didapatkan

dengan mendaftar pada twitter *developer* sebelumnya. Kode tersebut akan digunakan untuk mengakses data twitter melalui akun twitter yang dimiliki. Proses *collect* ini menggunakan keyword objek wisata seperti hastag atau kata yang berhubungan untuk mengambil data objek wisata pada twitter. Objek wisata pada tiap kabupaten/kota tersebut adalah Tabanan (#tanahlot, #batukaru, #Bedugul, #beratanlake), Karangasem (#besakih, #candidasa, #lempuyang), Gianyar (#monkeyforest, #sukawati, #ubud), Denpasar (#sanur, #kertalangu), Badung (#GWK, #Kuta, #puratamanayun, #uluwatu, #sangeh), Nusa Penida (#brokenbeach, #crystalbay, #kelingkingbeach). *Library* yang digunakan untuk *collect* data ini adalah tweepy yang mengambil data twitter kurang lebih selama 7-11 hari tergantung tanggal yang sudah di set pada program. Data twitter yang dikumpulkan adalah dalam bentuk JSON yang akan disimpan pada MONGODB *database*.

### **3.5.2 Filtering Data Twitter**

Proses Filtering merupakan tahap kedua dari proses analisis data twitter objek wisata. Penekanan pada proses ini adalah menyamakan atribut data json sehingga semua atribut pada data json yang telah disimpan bernilai sama, menghilangkan data json noise yang mengandung kata 'sale', 'jual', 'tanah', 'jokowigantipresiden', dll sehingga data memang data json yang berkaitan dengan objek wisata, dan mengambil atribut penting yang memang diperlukan untuk tahapan analisis.

### **3.5.3 Analisis data dengan Apache Spark**

Tahapan Analisis merupakan tahap ketiga dari proses analisis data twitter objek wisata. Proses analisis ini menggunakan apache spark sebagai tools. Proses analisis pertama terdiri dari menghitung banyaknya data *tweet* per objek wisata selama jangka waktu yang diperlukan (mis: sehari, seminggu, sebulan, dll) untuk mengetahui tingkat tren dari masing-masing objek wisata. Proses analisis kedua adalah menghitung jumlah favorited dan retweet dari masing-masing *tweet* untuk menunjang tingkat tren



masing-masing objek wisata. Data akhir yang diolah dari proses analisis akan diubah menjadi data dalam bentuk excel agar lebih mudah dalam proses visualisasi.

#### **3.5.4 Visualisasi Data**

Tahapan visualisasi data merupakan tahapan keempat dari proses analisis data twitter objek wisata. Tahapan ini merupakan tahapan untuk memvisualisasikan data excel yang diperoleh dari tahap analisis dalam bentuk grafik/diagram/peta/bar. Proses visualisasi ini akan menampilkan grafik data dari masing-masing objek wisata untuk mempermudah membaca tingkat popularitas dan tren dari masing-masing objek wisata.

## DAFTAR PUSTAKA

- Abed, A.A. and S.A. Rahman, *Python-based Raspberry Pi for Hand Gesture Recognition*.
- Anonim. 1990. Undang-Undang Republik Indonesia No. 9 Tahun 1990. Tentang Kepariwisataaan.Jakarta
- Chiang, O. 2011. Twitter hits nearly 200M accounts, 110M tweets per day, focuses on global expansion. Forbes.  
<<http://www.forbes.com/sites/oliverchiang/2011/01/19/twitter-hits-nearly-200m-users-110m-tweets-perday-focuses-on-global-expansion/>>.
- Chodorow, Kristina, 2013, *MongoDB The Definition Guide*, O'Reilly Media, Inc., 1005.
- Elzayady, Hossam, Badran, Khaled M., Salama, Gouda I., 2018 *Sentiment Analysis on Twitter Data using Apache Spark Framework*, IEEE.
- Hamstra, Mark, and Zaharia, Matei, 2013, *Spark: Cluster computing with working sets*, HotCloud, O'Reilly & Associates.
- Java, A., Song, X., Finin, T., & Tseng, B. 2007. *Why we twitter: Understanding microblogging usage and communities*. Paper presented at the 9<sup>th</sup> WebKDD and 1st SNA-KDD 2007 workshop on web mining and social network analysis, San Jose, California.
- Mahardhika,Yonathan S, Zuliarso, Eri, 2018, *Analisis Sentimen Terhadap Pemerintahan Joko Widodo Pada Media Sosial Twitter Menggunakan Algoritma Naives Bayes Classifier*, Prosiding SINTAK.
- Naaman, M., Boase, J., & Lai, C.-H. 2010. *Is it really about me? Message content in social awareness streams*. Paper presented at the 2010 ACM conference on computer supported cooperative work, Savannah,Georgia, USA.
- Oulasvirta, A., Lehtonen, E., Kurvinen, E., & Raento, M. 2010. *Making the ordinary visible in microblogs*. Personal Ubiquitous Computing, 14, 237–249.

Ranganathan, Jaishree, Irudayaraj, Allen S., Tzacheva, Angelina A, 2017, *Action Rules for Sentiment Analysis on Twitter Data using Spark*, IEEE.

Sakshi dan Upadhyaya, Shucitha, 2016, *Twitter Streaming API Using Apache Spark in Big Data Analytics*, International Journal of Scientific & Engineering Research,.

Twitter. Available online: <https://about.twitter.com/company>.

Apache Spark. Available online: <http://spark.apache.org>.