



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB III

METODOLOGI PENELITIAN

3.1. Objek Penelitian



Gambar 3.1. Logo Academy Awards

Sumber: (The Oscars, 2018)

Academy Awards yang biasa disebut Oscar, merupakan sebuah ajang internasional tahunan untuk menyatakan apresiasi terhadap karya film di bawah wawasan lembaga The Academy of Motion Picture Arts and Sciences (AMPAS). Diadakan semenjak tahun 1929, acara ini ditujukan tiap tahunnya kepada film, para pembuat film, aktor, dan aktris pada tahun sebelumnya, sehingga penghargaan-penghargaan ini membawa sebuah sejarah dan tradisi dunia perfilman yang panjang. Hingga sekarang, Academy Awards memiliki 24 kategori penghargaan dan sudah diikuti oleh ratusan negara di seluruh dunia sehingga terus menambahkan keragaman atas nilai budaya yang dibawa oleh setiap karyanya.

Penelitian ini merupakan percobaan pembuatan visualiasasi data dan prediksi yang akan ditampilkan dalam bentuk *dashboard* terhadap data film nominasi *Best Picture* dalam Academy Awards periode 1993 – 2017 yang didapatkan melalui kumpulan *dataset* yang berbeda untuk dieksplorasi.

Tabel 3.1. Sumber Pengambilan Dataset

No.	Asal Dataset	Nama Dokumen
1.	https://www.kaggle.com/rounakbanik/the-movies-dataset/	movies_metadata.csv keywords.csv
2.	https://datahub.io/rufuspollock/oscars-nominees-and-winners#data/	data_csv.csv
3.	https://datasets.imdbws.com/	rating data.tsv
4.	movies_metadata.csv	Genres.xlsx Prod Comps.xlsx
5.	keywords.csv	Keywords.xlsx
6.	data_csv.csv	Nominations.csv

Adapun dari keseluruhan parameter yang digunakan pada penelitian ini, hanya beberapa di antaranya yang akan diteliti lebih lanjut perihal peramalan prediksi yang akan terjadi di tahun 2018, yakni: *runtime*, *budget*, *revenue*, *rating*, *total vote*, dan *awards nominated*. Hal tersebut dikarenakan oleh faktor parameter lainnya yang tidak bersifat sebagai *continuous measure* sehingga tidak dapat diaplikasikan terhadap rumus peramalan yang sama, dimana pengertian *continuous measure* adalah suatu nilai dengan rentang tak terbatas yang setiap atributnya mewakili nilainya masing-masing.

3.2. Penelitian Terdahulu

Pada tabel 3.2. merupakan beberapa referensi jurnal penelitian terdahulu yang terdapat keterkaitan dengan penelitian ini.

Tabel 3.2. Penelitian Terdahulu

No.	Judul Jurnal	Pembahasan
1.	<p><i>Data Visualization of Poverty Level at Provinces in Indonesia from the year 2013-2015</i></p> <p>Peneliti: 1. Stevencua 2. Johan Setiawan</p> <p>Lokasi: <i>Department of Information Systems, Universitas Multimedia Nusantara, Tangerang, Indonesia</i></p> <p>Nama jurnal: IJNMT Vol. V, No. 1, Tahun 2018 ISSN: 2355-0082</p>	<p>Tujuan dari penelitian terdahulu ini adalah untuk membuat visualisasi data yang dapat menyampaikan suatu informasi mengenai distribusi kemiskinan tiap provinsi di Indonesia dan untuk menemukan hubungannya dengan tingkat pendidikan. Metode yang digunakan adalah visualisasi data <i>Eight-step</i> dan metodologi <i>data mining</i>. Data untuk penelitian ini diambil dari BPS dari tahun 2013 hingga 2015. Data yang dikumpulkan setelah <i>extraction</i> dan <i>cleansing</i> adalah sebanyak 6000 catatan. Waktu yang diperlukan untuk mengubah data dari beberapa publikasi dan memasukkannya ke dalam satu <i>file</i> Excel adalah 2 minggu. Tableau dipilih sebagai alat untuk membuat visualisasi data, dan ditampilkan dalam <i>dashboard story</i> yang terdiri dari 3 <i>dashboard</i>. <i>Dashboard</i> pertama adalah pemetaan data kemiskinan ke provinsi-provinsi di Indonesia. <i>Dashboard</i> kedua adalah pemetaan dan distribusi tingkat pendidikan di kota dan kabupaten di Indonesia. <i>Dashboard</i> ketiga menunjukkan peringkat daerah dengan tingkat persentase tertinggi warga negara pendidikan yang lebih rendah dari sekolah dasar di Indonesia. Berdasarkan Tes Penerimaan Pengguna kepada staf BPS, telah dapat menghasilkan <i>story</i> yang diperlukan dan interaktif. Kesimpulan dari visualisasi data ini adalah bahwa angka kemiskinan tinggi di daerah-daerah tertentu tidak langsung berbanding lurus dengan tingkat pendidikan.</p>

Tabel 3.2. Penelitian Terdahulu

<p>2.</p>	<p><i>Data Visualization Indicator Disease (Malaria, Dengue Fever, and Measles) in the Year 2012-2015</i></p> <p>Peneliti: 1. Immanuel Luigi Da Gusta 2. Johan Setiawan</p> <p>Lokasi: <i>Department of Information Systems, Universitas Multimedia Nusantara, Tangerang, Indonesia</i></p> <p>Nama jurnal: IJNMT Vol. IV, No. 2, Tahun 2017 ISSN: 2354-0082</p>	<p>Penelitian terdahulu ini ditujukan untuk membuat visualisasi data yang dapat membantu pemerintah dalam mengevaluasi perkembangan fasilitas kesehatan dalam hal sumber daya manusia untuk tenaga medis, membantu masyarakat untuk mengetahui jumlah distribusi rumah sakit dengan tenaga medis di wilayah regional, dan untuk memetakan indikator penyakit di Indonesia. Data untuk penelitian ini diambil dari BPS, dengan total 10.600 catatan setelah proses <i>Extract, Transform</i> dan <i>Load</i>. Waktu yang diperlukan untuk mengubah beberapa publikasi dari PDF, ke CSV dan kemudian ke <i>file Excel</i> adalah selama 3 minggu. Metode yang digunakan adalah visualisasi data <i>Eight-step</i> dan metodologi <i>data mining</i>. Tableau dipilih sebagai alat untuk membuat visualisasi data karena dapat menggabungkan setiap <i>dashboard</i> di dalam sebuah <i>story</i> interaktif, dan dapat memudahkan pengguna untuk menganalisis data. Hasilnya adalah <i>story</i> dengan 3 <i>dashboard</i> yang dapat memenuhi persyaratan dari staf BPS dan telah diuji dengan hasil yang memuaskan dalam <i>User Acceptance Test (UAT)</i>.</p>
<p>3.</p>	<p><i>Forecasting Data Centers Power Consumption with the Holt-Winters Method</i></p> <p>Peneliti: 1. Maurizio Rossi 2. Davide Brunelli</p> <p>Lokasi: <i>Department of Industrial Engineering, University of Trento, Trento, Italy</i></p>	<p>Dalam penelitian terdahulu ini ditunjukkan bagaimana alat peramalan dapat sangat meningkatkan efisiensi energi <i>data center</i>. Dilatarbelakangi oleh penggunaan <i>data center</i> yang telah meningkat pesat dalam beberapa tahun terakhir dan akan terus berlanjut di masa depan karena permintaan pasar untuk infrastruktur virtual, layanan <i>cloud</i>, dan aplikasi IoT. Dalam penelitian ini disajikan analisis dari metode peramalan yang sangat penting untuk menyelesaikan masalah</p>

Tabel 3.2. Penelitian Terdahulu

	<p>Nama jurnal: IEEE <i>Workshop</i> Pp. 210-214, Tahun 2015 ISBN: 978-1-4799-8215-8</p>	<p>optimasi. Penelitian ini memilih pendekatan referensi terhadap algoritma <i>exponential smoothing</i> Holt-Winters. Hasil akhirnya dapat mengungguli algoritma serupa yang diusulkan dalam literatur lainnya, dengan beban kerja yang sangat bervariasi dan sangat menarik bagi <i>data center</i> masa depan. Sebagai tahapan mendatang, direncanakan untuk membandingkan hasil ini dengan teknik peramalan lainnya berdasarkan pendekatan filter Kalman dan metode autoregresif seperti ARMA. Tujuan akhir adalah untuk menggambar seperangkat pedoman untuk peramalan energi yang efektif dari <i>data center</i> "hijau" generasi berikutnya.</p>
--	--	--

Kedua penelitian dari (Stevencua & Setiawan, 2018) dan (Da Gusta & Setiawan, 2017) yang terdapat pada tabel 3.2. merupakan penelitian yang dijadikan acuan dalam merancang visualisasi data berupa *dashboard* dengan menggunakan metode *Visual Data Mining* dengan menggunakan perangkat lunak Tableau *software tools*. Persamaan kedua jurnal penelitian tersebut adalah bahwa keduanya memakai perangkat lunak visualisasi data interaktif *business intelligence* yang sama, yakni Tableau *Software*. Metode yang digunakan untuk mengolah dan menampilkan *dataset* mentah yang diambil pun turut sama, walaupun tidak disebutkan secara lebih spesifik, yaitu *Visual Data Mining* (VDM) atau *eight-step data visualization* dengan metodologi *data mining*. Adapun perbedaan yang terdapat pada kedua jurnal penelitian tersebut dengan penelitian yang sedang dilakukan adalah objek *dataset* yang diteliti.

Diiringi dengan *Visual Data Mining*, metode pada penelitian ini juga ditambah dengan adanya *Exploratory Data Analysis (EDA)*. Kedua metode ini digunakan untuk mempermudah pengolahan data sehingga dapat divisualisasikan dan dapat dimengerti dengan mudah oleh pengguna awam. Sedangkan penelitian ini juga turut menambahkan fitur prediksi dengan menggunakan Holt-Winters *exponential smoothing* sebagai pengembangan dari penelitian terdahulu, dimana data prediksi ini secara langsung dapat diperoleh dari data visualisasi yang dilakukan sebelumnya.

3.3. *Visual Data Mining (VDM)*

Visual Data Mining merupakan metode utama yang akan digunakan dalam penelitian ini, yang dimana terdapat 3 tahapan utama, yaitu:

3.3.1. *Project Planning Phase*

Tujuan dari fase ini adalah untuk menemukan tujuan akhir dari penelitian yang akan dilakukan beserta dengan *scope* dari tujuan tersebut. Di dalam penelitian ini, *scope* yang digunakan adalah tipe *pilot* yang dapat membantu menginvestigasi, menganalisis, dan menjawab pertanyaan bisnis yang terdapat, namun tidak sampai melakukan *action plan*.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

3.3.2. *Data Preparation Phase*

Pada tahap ini, dilakukan pengumpulan data yang diambil dari berbagai kumpulan *dataset* yang berbeda, pemilihan, dan pengumpulan data sehingga tersimpan menjadi satu *dataset* yang utuh. Setelah itu, *dataset* ditransformasi dengan dilakukan proses *cleansing* agar tidak ada data yang *error*, *null*, atau inkonsisten, dan juga proses *filtering* agar format *dataset* yang nantinya diolah sudah sesuai tanpa adanya kekurangan/kelebihan data. Pada langkah terakhir dalam tahap ini, dilakukan proses pemeriksaan (*cross check*) terhadap *dataset* untuk memastikan apakah kedua proses *cleansing* dan *filtering* sebelumnya sudah berjalan sebagaimana mestinya, sehingga dapat mempermudah proses analisis data selanjutnya.

3.3.3. *Data Analysis Phase*

Pada tahapan terakhir ini, dilakukan proses seleksi *tools* yang akan digunakan dalam perancangan visualisasi data beserta dengan prediksinya, dimana *tools* yang akan digunakan tersebut adalah perangkat lunak Tableau. Selanjutnya, dilakukan eksplorasi data dengan menggunakan metode *Exploratory Data Analysis* (EDA) untuk mengetahui pola beserta tren data dan menerapkan rumus aritmatika sederhana untuk mempermudah proses pengolahan data setelahnya. Setelah semua data sudah siap untuk dipakai dan ditransformasi menjadi sebuah visualisasi interaktif yang berguna, dibangun keseluruhan *dashboard* beserta dengan semua isi konten *worksheet*nya dengan menerapkan metode *Visual Data Mining* (VDM).

3.4. *Exploratory Data Analysis (EDA)*

Exploratory Data Analysis (EDA) merupakan metode yang digunakan untuk membantu mempermudah pengolahan data melalui metode *Visual Data Mining (VDM)* dengan cara melakukan eksplorasi *dataset* dan menetapkan suatu rumus aritmatika sederhana dalam meringkas pengamatan yang ada. Metode ini dapat digunakan untuk melakukan eksplorasi secara menyeluruh terhadap *dataset* dengan cara membandingkan rata-rata, agregasi total, nilai minimum, nilai maksimum, dan lain-lain.

3.5. **Metode Penelitian**

Berikut merupakan keunggulan metode visualisasi atau khususnya analitik visual yang telah dijabarkan menurut (Ribarsky, Wang, & Dou, 2014).

1. Berdasarkan sifat alaminya, analitik visual mendukung eksplorasi, penemuan, dan pemikiran kompleks tentang data dan masalah.
2. Statistik, penggalian data, *machine learning*, pemrosesan sinyal, dan metode analitik mendalam lainnya terintegrasi dengan erat dengan visualisasi interaktif.
3. Bertujuan untuk menempatkan peran manusia pada titik yang tepat untuk menemukan wawasan utama, mengembangkan pemahaman yang mendalam, membuat keputusan, dan mengambil tindakan yang efektif.
4. Baik di sisi visualisasi dan analitik, analitik visual diposisikan secara efektif untuk mendukung pemahaman data berskala besar.

Metode final yang akan dipakai sesuai dengan penjelasan sebelumnya adalah metode *Visual Data Mining* (VDM) untuk menimalisir upaya yang digunakan dalam menafsir hasil dari visualisasi data yang akan dikemukakan sehingga turut dapat mengurangi kemungkinan adanya miskomunikasi dalam menginterpretasikan data.

3.6. Pemilihan *Tools* Visualisasi

Tabel 3.3. adalah perbandingan dari *tools* Tableau *software* dengan perangkat lunak *business intelligence* tradisional lainnya menurut (Zikri, Adrian, Soniawan, Azim, Dinur, & Akbar, 2017).

Tabel 3.3. — Perbandingan *Tools* Visualisasi Data

Tableau	<i>BI Tools</i> Tradisional
Tidak membutuhkan keahlian <i>programming</i> .	Mebutuhkan keahlian <i>programming</i> yang khusus.
Dapat mengkombinasikan berbagai jenis sumber data yang berbeda-beda, seperti <i>spreadsheet</i> , <i>database</i> , <i>cloud data</i> , dan <i>big data</i> .	Fokus hanya pada satu tipe basis data.
Menghemat waktu.	Memakan waktu.
Pengambil keputusan dapat secara langsung menggunakan <i>dashboard</i> untuk memperoleh informasi.	Pengambil keputusan harus meminta ahli IT untuk memperoleh informasi dari <i>database</i> .
<i>Query</i> dijalankan di balik layar.	Sangat tergantung pada bahasa <i>query</i> .
Perbedaan sumber data dapat dikombinasikan dengan mudah.	Mengkombinasi sumber data yang berbeda sulit untuk dilakukan.
Harga terjangkau.	Relatif mahal.
Solusi BI yang sempurna untuk bisnis yang kecil, medium dan luas.	Banyak dirancang untuk bisnis yang besar.

MULTIMEDIA
NUSANTARA

Tools yang akan dipakai sesuai dengan tabel 3.3. adalah perangkat lunak Tableau, dikarenakan juga oleh banyaknya keunggulan yang ditawarkan kepada para penggunanya. Salah satunya adalah kemudahan dalam melakukan proses *filtering* data. Selain itu, Tableau juga menawarkan pilihan visualisasi data yang cukup beragam dan dapat disesuaikan dengan keinginan pengguna, tanpa menambah kesulitan dalam penggunaannya. Ditambah dengan komunitas pengguna yang besar, *e-learning* informatif yang memadai, dan kisaran harganya yang terjangkau, membuat Tableau *software* menjadi pilihan yang tepat bagi seluruh kalangan masyarakat pemakai *business intelligence*.

3.7. Metode Prediksi

Berikut merupakan keunggulan algoritma prediksi Holt-Winters dibandingkan dengan pendekatan tradisional lainnya yang telah dijabarkan menurut (Rossi & Brunelli, 2015).

1. Telah ditunjukkan sebagai solusi optimal untuk kelas yang sangat umum dari *state-space model* dengan persyaratan memori yang lebih rendah daripada metode *auto-regressive* yang lebih umum.
2. Hasil yang dicapai menunjukkan kinerja yang lebih baik dengan beban kerja yang sangat bervariasi dan pada rincian waktu yang diusulkan.
3. Sudah banyak penelitian yang dikerjakan terhadap skema ini dan terdapat banyak versi algoritma yang dirancang untuk skenario tertentu.
4. Jika dibandingkan dengan pemulusan eksponensial standar (ARMA, ARIMA, dan sebagainya) adalah berkurangnya kebutuhan memori dan dapat disesuaikannya fase pelatihan secara otomatis.