



## BAB II

### LANDASAN TEORI

© Hak cipta dimiliki IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

#### A. Tinjauan Pustaka

Hak Cipta Dilindungi Undang-Undang

##### 1. Informasi

Menurut Kenneth C. Laudon dan Jane P. Laudon (2018 : 44) definisi informasi yaitu :

*“ By information we mean data that have been shaped into a form that is meaningful an useful to human beings.”*

Informasi dimaksud dengan data yang telah terbentuk menjadi sebuah yang sangat berarti bagi manusia.

##### 2. Sistem Informasi

Menurut Kenneth C. Laudon dan Jane P. Laudon (2018 : 44) definisi Sistem Informasi yaitu :

*“An information system can be defined technically as a set of interrelated components that collect (or retrieve), process, store, and distribute information to support decision making and control in an organization.”*

Sistem informasi dapat didefinisikan secara teknis sebagai sekumpulan komponen yang saling terkait yang mengumpulkan (atau mendapatkan kembali), memproses, menyimpan, dan mendistribusikan informasi untuk mendukung pengambilan keputusan dan pengendalian dalam suatu organisasi.

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



### 3. Data

Menurut Pang-Ning Tan (2019:43) definisi data yaitu :

*“Data sets differ in a number of ways. For example, the attributes used to describe data objects can be of different types—quantitative or qualitative—and data sets often have special characteristics; e.g., some data sets contain time series or objects with explicit relationships to one another. Not surprisingly, the type of data determines which tools and techniques can be used to analyze the data. Indeed, new research in data mining is often driven by the need to accommodate new application areas and their new types of data.”*

Kumpulan data berbeda dalam beberapa cara. Misalnya, atribut yang digunakan untuk mendeskripsikan objek data bisa dari tipe yang berbeda—kuantitatif atau kualitatif—dan rangkaian data seringkali memiliki karakteristik khusus; misalnya, beberapa data set berisi deret waktu atau objek dengan hubungan eksplisit satu sama lain. Tidak mengherankan, jenis data menentukan alat dan teknik mana yang bisa akan digunakan untuk menganalisis data. Memang, penelitian baru dalam penambangan data sering kali didorong oleh kebutuhan untuk mengakomodasi area aplikasi baru dan tipe barunya data..

Menurut Pang-Ning Tan (2019:43) kualitas data yaitu :

*“Data is often far from perfect. While most data mining techniques can tolerate some level of imperfection in the data, a focus on understanding and improving data quality typically improves the quality of the resulting analysis. Data quality issues that often need to be addressed include the presence of noise and outliers; missing, inconsistent, or duplicate data; and data that is biased or, in some other way, unrepresentative of the phenomenon or population that the data is supposed to describe.”*

Data seringkali jauh dari sempurna. Sementara sebagian besar data teknik penambangan dapat mentolerir beberapa tingkat ketidaksempurnaan dalam data, sebuah fokus pada pemahaman dan peningkatan kualitas data biasanya

Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



meningkatkan kualitas dari analisis yang dihasilkan. Masalah kualitas data yang sering perlu ditangani termasuk adanya kebisingan dan outlier; hilang, tidak konsisten, atau duplikat data; dan data yang bias atau, dengan cara lain, tidak mewakili fenomena atau populasi yang seharusnya dideskripsikan oleh data tersebut.

Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

Menurut Pang-Ning Tan (2019:46) Tipe data ada dua, yaitu :

Kumpulan data seringkali dapat dilihat sebagai kumpulan objek data. Nama lain untuk objek data adalah record, point, vector, pattern, event, case, sample, instance, observasi, atau entitas.

Pada gilirannya, objek data dijelaskan oleh sejumlah atribut yang menangkap karakteristik suatu objek, seperti massa objek fisik atau waktu di mana suatu peristiwa terjadi. Nama lain untuk an atribut adalah variabel, karakteristik, bidang, fitur, atau dimensi.

Menurut Pang-Ning Tan (2019:54) Karakteristik data ada dua, yaitu :

**Dimensi** Dimensi dimensi kumpulan data adalah jumlah atribut yang dimiliki objek dalam kumpulan data. Menganalisa data dengan asejumlah kecil dimensi cenderung berbeda secara kualitatif dari analisis data berdimensi sedang atau tinggi.

**Distribusi** Distribusi suatu kumpulan data adalah frekuensi kejadian dari berbagai nilai atau kumpulan nilai untuk atribut yang terdiri dari objek data. Secara ekuivalen, distribusi kumpulan data dapat dianggap sebagai deskripsi konsentrasi objek di berbagai wilayah ruang data

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



#### 4. Data Mining

Menurut Charu C. Aggarwal (2015 : 1) definisi Data Mining yaitu :

*“Data mining is the study of collecting, cleaning, processing, analyzing, and gaining useful insights from data. A wide variation exists in terms of the problem domains, applications, formulations, and data representations that are encountered in real applications. Therefore, “data mining” is a broad umbrella term that is used to describe these different aspects of data processing.”*

Penambangan data adalah studi tentang mengumpulkan, membersihkan, memproses, menganalisis, dan memperoleh manfaat wawasan dari data. Variasi yang luas ada dalam hal domain masalah, aplikasi , formulasi, dan representasi data yang ditemui dalam aplikasi nyata. Karena itu, "penambangan data" adalah istilah umum yang digunakan untuk menggambarkan berbagai aspek ini pengolahan data.

#### 5. Time Series Forecasting

Menurut Charu C. Aggarwal (2015 : 464 ) definisi *Time Series Forecasting* yaitu :

*“Forecasting is one of the most common applications of time series analysis. The prediction of future trends has applications in retail sales, economic indicators, weather forecasting, stock markets, and many other application scenarios. In this case, we have one or more series of data values, and it is desirable to predict the future values of the series using the history of previous values.”*

Peramalan adalah salah satu aplikasi analisis deret waktu yang paling umum. Prediksi trend masa depan memiliki aplikasi dalam penjualan ritel, indikator ekonomi, prakiraan cuaca, pasar saham, dan banyak skenario aplikasi lainnya. Dalam hal ini, kami memiliki satu atau lebih rangkaian nilai data, dan diinginkan untuk memprediksi nilai masa depan dari rangkaian tersebut menggunakan riwayat nilai sebelumnya.

Hak Cipta Dilindungi Undang-Undang

© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



## 6. ARIMA Model

Menurut Robert H. Shumway dan David S. Stoffer (2017 : 75) definisi dari *Autoregressive Models* yaitu :

*“Autoregressive models are based on the idea that the current value of the series,  $x_t$ , can be explained as a function of  $p$  past values,  $x_{t-1}, x_{t-2}, \dots, x_{t-p}$ , where  $p$  determine sthe number of steps into the past needed to forecast the current value.”*

Model autoregresif didasarkan pada gagasan bahwa nilai saat ini dari deret,  $x_t$ , dapat dijelaskan sebagai fungsi dari  $p$  nilai lampau,  $x_{t-1}, x_{t-2}, \dots, x_{t-p}$ , di mana  $p$  menentukan jumlah langkah ke masa lalu diperlukan untuk meramalkan nilai saat ini.

Menurut Robert H. Shumway dan David S. Stoffer (2017 : 77) contoh model

*Autoregressive* yaitu :

Kami memulai penyelidikan model AR dengan mempertimbangkan model orde pertama, AR(1), diberikan oleh  $x_t = \phi x_{t-1} + w_t$ . Iterasi mundur  $k$  waktu, kita dapatkan

Metode ini menunjukkan bahwa, dengan terus mengulang ke belakang, dan asalkan  $|\phi| < 1$  dan  $\text{supt var}(x_t) < \infty$ , kita dapat merepresentasikan model AR(1) sebagai proses linier diberikan oleh 1

Menurut Robert H. Shumway dan David S. Stoffer (2017 : 75) definisi dari

*Moving Average Models* yaitu :

*“As an alternative to the autoregressive representation in which the  $x_t$  on the left-hand side of the equation are assumed to be combined linearly, the moving average model of order  $q$ , abbreviated as MA( $q$ ), assumes the white noise  $w_t$  on the right-hand side of the defining equation are combined linearly to form the observed data.”*

Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



Sebagai alternatif representasi autoregresif di mana  $x_t$  di sebelah kiri sisi persamaan diasumsikan digabungkan secara linier, model rata-rata bergerak orde  $q$ , disingkat MA( $q$ ), mengasumsikan white noise  $w_t$  di sisi kanan dari persamaan pendefinisian digabungkan secara linear untuk membentuk data yang diamati.

Hak Cipta Dilindungi Undang-Undang

© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Institut Bisnis dan Informatika Kwik Kian Gie

Menurut Robert H. Shumway dan David S. Stoffer (2017 : 75) operator dari *Moving Average Models* yaitu :

Pembaca dapat memverifikasi bahwa ini stasioner dan berbentuk AR(1)  $x_t = \phi x_{t-1} + w_t$ . Sayangnya, model ini tidak berguna karena mengharuskan kita mengetahui masa depan yang akan datang mampu memprediksi masa depan. Ketika suatu proses tidak bergantung pada masa depan, seperti AR(1) ketika  $|\phi| < 1$ , kita akan mengatakan prosesnya kausal. Dalam kasus ledakan dari contoh ini, prosesnya stasioner, tetapi juga bergantung pada masa depan, dan tidak kausal.

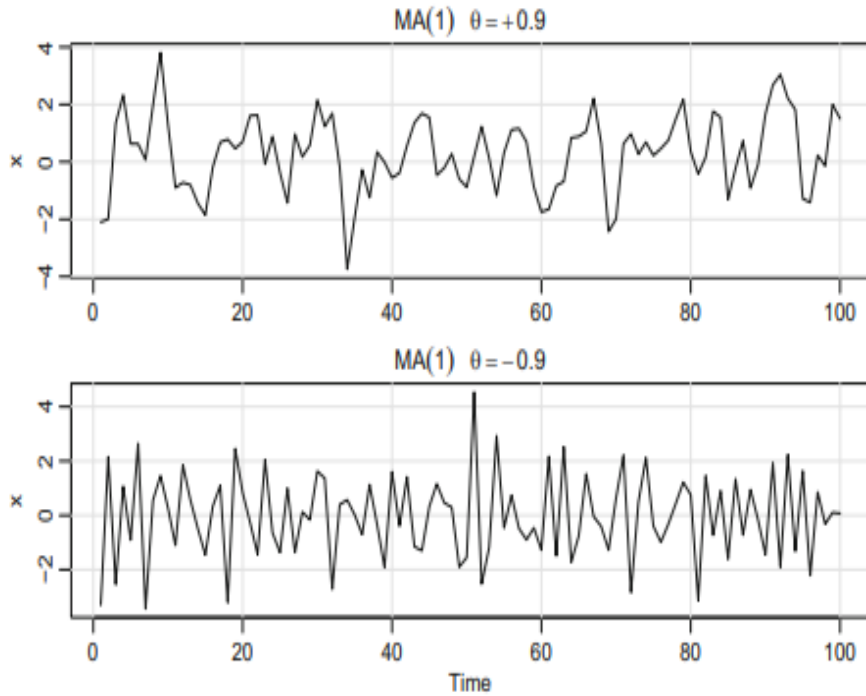
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



**Gambar 2. 1**

Simulated MA(1) models

(Sumber : Menurut Robert H. Shumway dan David S. Stoffer (2017 : 775))

Menurut Robert H. Shumway dan David S. Stoffer (2017 : 75) langkah

membuat ARIMA model yaitu :

*“There are a few basic steps to fitting ARIMA models to time series data. These steps involve • plotting the data, • possibly transforming the data, • identifying the dependence orders of the model, • parameter estimation, • diagnostics, and • model choice.”*

Ada beberapa langkah dasar untuk menyesuaikan model ARIMA dengan data deret waktu. Langkah-langkah ini meliputi • memplot data, • kemungkinan mengubah data, • mengidentifikasi urutan ketergantungan model, • estimasi parameter, • diagnostik, dan • pemilihan model.



## 7. Blockchain

Menurut Daniel Drescher (2017 : 17) fungsi *Blockchain* yaitu :

*“The blockchain is a tool for achieving integrity in distributed software systems. Hence, it can be seen as a tool to achieve a nonfunctional aspect of the implementation layer.”*

Blockchain adalah alat untuk mencapai integritas dalam sistem perangkat lunak terdistribusi. Oleh karena itu, dapat dilihat sebagai alat untuk mencapai aspek nonfungsional dari lapisan implementasi.

Layer	Functional Aspects	Nonfunctional Aspects
Application	Deposit money Withdraw money Transfer money Monitor account balance	The graphical user interface looks beautiful Easy to use Transfer of money is done fast System has many participants
Implementation	?	Available 24 hours a day Fraud resistant Maintaining integrity Ensure user privacy

**Gambar 2. 2**  
Fungsi Blockchain  
(Sumber : Daniel Drescher (2017 : 10))

Menurut Daniel Drescher (2017 : 17) definisi *peer-to-peer* yaitu :

*“ Peer-to-peer systems are distributed software systems that consist of nodes (individual computers), which make their computational resources (e.g., processing power, storage capacity, or information distribution) directly available to another. When joining a peer-to-peer system, users turn their computers into nodes of the system that are equal concerning their rights and roles. Although users may differ with respect to the resources they contribute, all the nodes in the system have the same functional capability and responsibility. Hence, the computers of all users are both suppliers and consumers of resources.*

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.





*For example, in a peer-to-peer file sharing system, the individual files are stored on the users' machines. When someone wants to download a file in such a system, he or she is downloading it from another person's machine, which could be the next door neighbor or someone located halfway around the world."*

Sistem peer-to-peer adalah sistem perangkat lunak terdistribusi yang terdiri dari node (komputer individu), yang membuat sumber daya komputasi mereka (misalnya, kekuatan pemrosesan, kapasitas penyimpanan, atau distribusi informasi) tersedia secara langsung ke yang lain. Saat bergabung dengan sistem peer-to-peer, pengguna menghidupkan komputer mereka ke dalam simpul-simpul sistem yang sama mengenai hak dan peran mereka. Meskipun pengguna mungkin berbeda sehubungan dengan sumber daya yang mereka sumbangkan, semua node dalam sistem memiliki kemampuan fungsional dan tanggung jawab yang sama. Oleh karena itu, komputer dari semua pengguna adalah pemasok dan konsumen sumber daya.

Misalnya, dalam sistem berbagi file peer-to-peer, masing-masing file disimpan di mesin pengguna. Ketika seseorang ingin mengunduh file di sistem seperti itu, dia mengunduhnya dari mesin orang lain, yang bisa jadi tetangga sebelah atau seseorang yang terletak di tengah jalan Dunia.

## 8. Cryptocurrency

Menurut Antony Lewis (2018 : 241) definisi Ethereum yaitu :

*"The vision of Ethereum is to create an unstoppable, censorship resistant, self-sustaining, decentralised, world computer. To achieve this, Ethereum builds on the concepts we saw with Bitcoin. If you consider Bitcoin as trustless validation and distributed storage of (transaction) data, Ethereum is trustless validation and distributed storage and processing of data and logic."*

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



Visi Ethereum adalah untuk menciptakan yang tak terhentikan, tahan sensor,

mandiri, terdesentralisasi, komputer dunia. Untuk mencapai ini, Ethereum dibangun di atas konsep yang kami lihat dengan Bitcoin. Jika Anda menganggap Bitcoin sebagai validasi tanpa kepercayaan dan penyimpanan data (transaksi) terdistribusi, Ethereum adalah validasi tanpa kepercayaan dan penyimpanan dan pemrosesan terdistribusi data dan logika.

Menurut Antony Lewis (2018 : 117) definisi *Cryptography* yaitu :

*“Cryptography is, however, not just for spies, criminals, and terrorists. It is now used extensively to protect data that travels across the internet. The ‘s’ in ‘https’ stands for secure. It means that cryptography is being used to guarantee that the website you think you are visiting is in fact the genuine website. It also means that the data in flight between you and that website is encrypted or jumbled up, so snoopers can’t easily read the communications between your device and the website that you are accessing.”*

Kriptografi, bagaimanapun, tidak hanya untuk mata-mata, penjahat, dan teroris. Dia sekarang digunakan secara luas untuk melindungi data yang melintasi internet. Itu 's' di 'https' berarti aman. Artinya kriptografi sedang digunakan menjamin bahwa situs web yang menurut Anda Anda kunjungi sebenarnya adalah asli situs web. Ini juga berarti bahwa data dalam penerbangan antara Anda dan itu situs web dienkripsi atau dikacaukan, sehingga pengintai tidak dapat dengan mudah membaca komunikasi antara perangkat Anda dan situs web tempat Anda berada mengakses.

## 9. Python

Menurut John Paul Mueller (2018 : 1) definisi Python yaitu :

*“Python is an example of a language that does everything right within the domain of things that it’s designed to do.”*



Python adalah contoh bahasa yang melakukan segalanya dengan benar dalam domain hal-hal yang dirancang untuk dilakukannya.

Menurut John Paul Mueller (2018 : 1) cara kerja Python yaitu :

*“Python emphasizes code readability and a concise syntax that lets you write applications using fewer lines of code than other programming languages require. You can also use a coding style that meets your needs, given that Python supports the functional, imperative, object-oriented, and procedural coding styles.”*

Python menekankan keterbacaan kode dan sintaks ringkas yang memungkinkan Anda menulis aplikasi yang menggunakan lebih sedikit baris kode daripada yang dibutuhkan bahasa pemrograman lain. Anda juga dapat menggunakan gaya pengkodean yang memenuhi kebutuhan Anda, mengingat dukungan Python gaya pengkodean fungsional, imperatif, berorientasi objek, dan prosedural.

Menurut Wes McKinney (2018 : 2) fungsi Python untuk data analisis yaitu :

*“For data analysis and interactive computing and data visualization, Python will inevitably draw comparisons with other open source and commercial programming languages and tools in wide use, such as R, MATLAB, SAS, Stata, and others. In recent years, Python’s improved support for libraries (such as pandas and scikit-learn) has made it a popular choice for data analysis tasks. Combined with Python’s overall strength for general-purpose software engineering, it is an excellent option as a primary language for building data applications.”*

Untuk analisis data dan komputasi interaktif dan visualisasi data, Python pasti akan tabel menarik perbandingan dengan open source dan bahasa pemrograman komersial lainnya pengukur dan alat yang digunakan secara luas, seperti R, MATLAB, SAS, Stata, dan lain-lain. Baru-baru ini tahun, dukungan Python yang ditingkatkan untuk perpustakaan (seperti panda dan scikit-learn) telah

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



menjadikannya pilihan populer untuk tugas analisis data. Dikombinasikan dengan keseluruhan Python kekuatan untuk rekayasa perangkat lunak tujuan umum, itu adalah pilihan yang sangat baik sebagai prioritas bahasa marie untuk membangun aplikasi data.

Menurut Wes McKinney (2018 : 4) *Essential Python Libraries* yaitu :

*“NumPy, short for Numerical Python, has long been a cornerstone of numerical computing in Python. It provides the data structures, algorithms, and library glue needed for most scientific applications involving numerical data in Python.”*

*pandas provides high-level data structures and functions designed to make working with structured or tabular data fast, easy, and expressive. Since its emergence in 2010, it has helped enable Python to be a powerful and productive data analysis environment. The primary objects in pandas that will be used in this book are the DataFrame, a tabular, column-oriented data structure with both row and column labels, and the Series, a one-dimensional labeled array object.*

*matplotlib is the most popular Python library for producing plots and other two dimensional data visualizations. It was originally created by John D. Hunter and is now maintained by a large team of developers. It is designed for creating plots suitable for publication. While there are other visualization libraries available to Python programmers, matplotlib is the most widely used and as such has generally good integration with.”*

NumPy, kependekan dari Numerical Python, telah lama menjadi landasan komputasi numerik dengan Python. Ini menyediakan struktur data, algoritme, dan perekat perpustakaan yang diperlukan untuk sebagian besar aplikasi ilmiah yang melibatkan data numerik dengan Python.

Panda menyediakan struktur dan fungsi data tingkat tinggi yang dirancang untuk bekerja dengan data terstruktur atau tabular dengan cepat, mudah, dan ekspresif. Sejak kemunculannya pada tahun 2010, ini telah membantu Python menjadi lingkungan analisis data yang kuat dan produktif. Objek utama dalam panda yang akan digunakan dalam buku ini adalah DataFrame, sebuah struktur data

Hak Cipta Dilindungi Undang-Undang

Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



berorientasi kolom dengan label baris dan kolom, dan Seri, sebuah objek larik berlabel satu dimensi.

Matplotlib adalah pustaka Python paling populer untuk menghasilkan plot dan visualisasi data dua dimensi lainnya. Ini awalnya dibuat oleh John D. Hunter dan sekarang dikelola oleh tim pengembang yang besar. Ini dirancang untuk membuat plot yang cocok untuk publikasi. Meskipun ada perpustakaan visualisasi lain yang tersedia untuk pemrograman Python, matplotlib adalah yang paling banyak digunakan dan umumnya memiliki integrasi yang baik dengan.

## 10. Orange

Orange adalah sebuah platform data mining untuk menjalankan analisis data dan visualisasi. Tujuan dari pemanfaatan Orange adalah menyediakan platform model prediksi dan sistem rekomendasi. Contoh penggunaan Orange pada pembelajaran selain data mining adalah bioinformatika, penelitian genomic dan lain-lain. Orange menggunakan bahasa pemrograman C++, Python dan Cython.

## B. Penelitian Terdahulu

### 1. Prediksi Harga Cryptocurrency Dengan Metode *K-Nearest Neighbours*

Hasil penelitian yang dilakukan oleh Haerul Fatah dan Agus Subekti, bahwa kesimpulannya :

Dari hasil eksperimen metode KNN (*K Nearest Neighbors*) pada 3 buah dataset dengan menggunakan parameter nilai K dan Nearest neighbour Search Algorithm, dapat disimpulkan model KNN yang memiliki akurasi paling baik adalah



KNN dengan nilai  $K=3$  dan *Nearest Neighbour Search Algorithm*: Linear KNN

Search.

Semakin besar nilai  $K$  maka semakin besar pula nilai Mean Absolute Error dan nilai Root Mean Squared Error nya. Serta semakin banyak dataset/data history harga Cryptocurrency yang digunakan, maka semakin kecil nilai Mean Absolute Error dan nilai Root Mean Squared Error nya.

Penelitian yang dilakukan melakukan prediksi terhadap nilai Close harga Cryptocurrency pada waktu tertentu, berbeda dengan penelitian penelitian terkait yang memprediksi kondisi/arrah dari Cryptocurrency apakah naik atau turun.

## 2. Aplikasi Prediksi Jangka Pendek Harga Bitcoin Menggunakan Metode Arima

Hasil penelitian yang dilakukan oleh Nur Fitriani Bintang Pradana dan Sri Lestanti, bahwa kesimpulannya :

Penelitian ini yaitu ARIMA telah berhasil diterapkan pada aplikasi berbasis web untuk melakukan prediksi harga bitcoin untuk satu hingga tujuh hari kedepan dengan hasil yang baik. Beberapa model telah diuji dan model ARIMA (3,1,3) dapat melakukan prediksi harga bitcoin dengan tingkat akurasi yang baik. Rata-rata nilai MAPE yang dihasilkan adalah sebesar 0,84 dengan rentan nilai sebesar 1,34 untuk prediksi hari pertama (1 Juli 2020) dan 0,98 untuk prediksi hari ketujuh (7 Juli 2020). Performa ARIMA semakin baik jika digunakan untuk prediksi jangka pendek, terutama untuk prediksi periode dua hari kedepan. Semakin banyak periode yang diprediksi semakin rendah tingkat akurasinya. ARIMA dapat digunakan untuk prediksi harga bitcoin namun dengan selisih harga yang cukup signifikan tapi mampu memprediksi pola perubahan harga dalam periode tujuh hari kedepan.

Hak Cipta Dilindungi Undang-Undang

Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



Karakteristik perubahan harga bitcoin yang tidak menentu atau berpola membuat sulit untuk menentukan model ARIMA yang sesuai. Satu model yang baik saat melakukan prediksi pada satu rentang periode belum tentu baik untuk rentang periode berikutnya.

Adapun saran untuk mengembangkan penelitian ini yaitu menggunakan dataset dengan frekuensi perubahan data yang lebih cepat, seperti data perubahan bitcoin dalam hitungan menit. Mengkombinasikan metode ARIMA dengan metode lainnya seperti sentiment analysis dan jaringan syaraf agar dapat menghasilkan tingkat akurasi yang lebih baik lagi.

### 3. Implementasi Orange Data Mining Untuk Prediksi Harga Bitcoin

Hasil penelitian yang dilakukan oleh Indriyanti , Nurul Ichsan , Haerul Fatah , Tri Wahyuni , Erni Ermawati , bahwa kesimpulanya :

Hasil implementasi orange data mining untuk prediksi harga bitcoin dapat dilihat pada widget Test and Score, dimana diperoleh hasil nilai RMSE dan MAE dari masing-masing model. Nilai RMSE dan MAE dari metode K-Nearest Neighbor (K-NN) sebesar 0.036 dan 0.027. Sedangkan metode Support Vector Machine memiliki nilai RMSE dan MAE sebesar 0.010 dan 0.008.

Berdasarkan nilai RMSE dan MAE tersebut dapat disimpulkan bahwa metode Support vector Machine (SVM) merupakan metode terbaik untuk memprediksi perubahan harga bitcoin dibandingkan dengan metode K-Nearest Neighbor (K-NN). Metode terbaik pada penelitian ini dapat diteliti kembali untuk penelitian yang akan mendatang. Penelitian selanjutnya disarankan dapat

Hak Cipta Dilindungi Undang-Undang

Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.





membandingkan metode SVM dengan metode lainnya untuk memprediksi harga bitcoin menggunakan tools orange atau dapat menerapkan tools lainnya.



Hak cipta milik IBIKKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

#### 4. Prediksi Harga Bitcoin Menggunakan Metode Random Forest

Hasil penelitian yang dilakukan oleh Siti Saadah , Haifa Salsabila bahwa kesimpulannya :

Berdasarkan hasil penelitian, algoritma random forest dengan menggunakan atribut low, high dan price berhasil memprediksi harga bitcoin dengan performansi yang baik dengan nilai MAPE sebesar 1.50% atau dengan akurasi yang diperoleh sekitar 98% dengan menggunakan data acak. Perolehan akurasi untuk prediksi dengan data acak ini menunjukkan nilai yang setara yakni sekitar 95% - 98%. Sehingga mengacu kepada hasil prediksi ini dapat disimpulkan bahwa untuk memprediksi data yang memiliki karakteristik seperti bitcoin, yakni memiliki fluktuasi tinggi, random forest dikatakan berhasil memberikan fitting yang sesuai dengan data sesungguhnya. Akan tetapi, hasil ini berbanding terbalik pada saat memprediksi menggunakan data tidak acak.

Keberhasilan pada penelitian ini dapat menjadi pilot project untuk penelitian selanjutnya mengenai penelitian yang terkait dengan pergerakan nilai mata uang digital, seperti pada saat mengamati prediksi pergerakan global financial risk. Mengingat bahwa pembelajaran mengenai prediksi mata uang digital juga dapat memberikan efek untuk stabilitas nilai keuangan di sebuah negara secara tidak langsung, oleh sebab itu tidak menutup kemungkinan diperlukan penelitian yang

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



menggabungkan algoritma ini dengan metode deep learning di masa yang akan datang.

Hak cipta milik IBIKKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

## 5. Peramalan Cryptocurrency Dengan Autoregressive Integrated Moving Average (ARIMA) Dan Risiko Kerugian Dengan Value At Risk (VAR)

Hasil penelitian yang dilakukan oleh Amalia Listiani , Dana AL Mahkya

bahawa kesimpulannya :

Penelitian ini menggunakan model Autoregressive Integrated Moving

Average (ARIMA) dalam peramalan harga Bitcoin (BTC) dalam 10 hari kedepan.

Hasil penelitian menunjukkan bahwa model ARIMA [4,1,2] baik dalam melakukan

prediksi BTC dengan MAPE sebesar 6%. Berdasarkan hasil penelitian dengan Value

at Risk (VaR) seorang atau perusahaan akan mengalami kerugian maksimum sebesar

5,85% jika berinvestasi di BTC dan terdapat 5% kemungkinan akan melebihi 5,85%

kerugian yang akan dialami. Diharapkan penelitian ini dapat menjadi pertimbangan

dalam melakukan investasi kedepannya.

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.