

PREDIKSI HASIL PERTANDINGAN LIGA BOLA BASKET NBA MUSIM 2022/2023 DENGAN MENGUNAKAN METODE *QUADRATIC MULTIVARIATE*

LAPORAN TUGAS AKHIR

Oleh:

Boy Cakraningrat

105217028



**FAKULTAS SAINS DAN ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER
UNIVERSITAS PERTAMINA
FEBRUARI 2023**

1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah kecependekian;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

PREDIKSI HASIL PERTANDINGAN LIGA BOLA BASKET NBA MUSIM 2022/2023 DENGAN MENGUNAKAN METODE *QUADRATIC MULTIVARIATE*

LAPORAN TUGAS AKHIR

Oleh:

Boy Cakraningrat

105217028



**FAKULTAS SAINS DAN ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER
UNIVERSITAS PERTAMINA
FEBRUARI 2023**

1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah kecependekian;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.



Universitas
Pertamina

Hak Cipta milik Universitas Pertamina

Dilindungi Undang-Undang

© Copyright of Universitas Pertamina



1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah kecependekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

Universitas
Pertamina

LEMBAR PENGESAHAN

Universitas
Pertamina



Judul Tugas Akhir

: PREDIKSI HASIL PERTANDINGAN LIGA
BOLA BASKET NBA MUSIM 2022/2023 DEN-
GAN MENGGUNAKAN METODE *Quadratic*
Multivariate

Nama Mahasiswa

: Boy Cakraningrat

Nomor Induk Mahasiswa

: 105217028

Program Studi

: Ilmu Komputer

Fakultas

: Sains dan Ilmu Komputer

Tanggal Lulus Sidang Tugas Akhir

: 15 februari 2023

Jakarta, 14 MARET 2023

MENGESAHKAN

Pembimbing I

Rangga Ganzar Noegraha, Ph.D
NIP.119005

Pembimbing II

Ade Irawan, Ph.D
NIP.116130

MENGETAHUI,

Ketua Program Studi

Muhamad Koyimatu, Ph.D
NIP. 116153

© Copyright of Universitas Pertamina

Hak Cipta Universitas Pertamina

Undang-Undang-Undang

1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah keceandekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa Tugas Akhir berjudul PREDIKSI HASIL PERTANDINGAN LIGA BOLA BASKET NBA MUSIM 2022/2023 DENGAN MENGGUNAKAN METODE *Quadratic Multivariate* ini adalah benar-benar merupakan hasil karya saya sendiri dan tidak mengandung materi yang ditulis oleh orang lain kecuali telah dikutip sebagai referensi yang sumbernya telah dituliskan secara jelas sesuai dengan kaidah penulisan karya ilmiah.

Apabila dikemudian hari ditemukan adanya kecurangan dalam karya ini, saya bersedia menerima sanksi dari Universitas Pertamina sesuai dengan peraturan yang berlaku.

Demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Pertamina hak bebas royalti noneksklusif (*non-exclusive royalty-free right*) atas Tugas Akhir ini beserta perangkat yang ada. Dengan hak bebas royalti noneksklusif ini Universitas Pertamina berhak menyimpan, mengalih media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya

Jakarta, 14 MARET 2023

Yang membuat pernyataan,



Boy Cakraningrat

ABSTRAK

Boy Cakraningrat. 105217028. PREDIKSI HASIL PERTANDINGAN LIGA BOLA BASKET NBA MUSIM 2022/2023 DENGAN MENGGUNAKAN METODE *Quadratic Multivariate* .

Memprediksi hasil pertandingan bola basket di NBA merupakan hal yang penting bagi para penggemar dan staf manajemen tim. Tujuan dari penelitian ini adalah untuk memprediksi hasil pertandingan bola basket NBA menggunakan metode polinomial derajat dua multivariat dan membandingkannya dengan regresi linier. Dalam pembuatan model prediksi ini terdapat dua jenis dataset yang digunakan untuk menganalisis pengaruh poin yang dicetak pemain, tembakan 3 poin yang berhasil dilakukan pemain, tembakan bebas yang berhasil dilakukan pemain, dan umpan bola terakhir dari pemain dalam satu tim sebelum mencetak poin, terhadap total poin timnya. Dataset tersebut dinamakan Dataset 18 dan Dataset 22. Dataset 18 digunakan untuk membangun model pertama yang disebut Model 18, dan Dataset 22 digunakan untuk membangun model kedua yang disebut Model 22. Hasil penelitian menunjukkan bahwa metode polinomial *multivariate quadratic* pada Model 22 memberikan hasil prediksi dengan akurasi yang lebih tinggi dibandingkan regresi linier, yaitu sebesar 0.84278. Hasil ini dapat digunakan sebagai bantuan bagi penggemar dan staf manajemen tim untuk memprediksi hasil pertandingan bola basket di NBA.

Kata kunci: Pertandingan Nba, *Machine learning*, korelasi Pearson, evaluasi metrix, regresi *multivariate*, regresi *linear*.

ABSTRACT

Boy Cakraningrat. 105217028. PREDICTION OF BASKETBALL LEAGUE MATCH RESULTS IN NBA USING multivariate Quadratic METHOD.

Predicting the results of NBA basketball games is important for both fans and team management staff. The purpose of this study is to predict the outcomes of NBA basketball games using a multivariate quadratic polynomial method and compare it with standard linear regression. Two types of datasets were used in creating this prediction model to analyze how a team's total points are influenced by player scoring, successful 3-point shooting by players, successful free throws by players, and the last pass by a player before scoring. These datasets are named Dataset 18 and Dataset 22. Dataset 18 was used to build the first model, called Model 18, and Dataset 22 was used to build the second model, called Model 22. The results of the study indicate that the multivariate quadratic polynomial method used in Model 22 provided more accurate predictions, with an accuracy of 0.84278. These findings can be used as a tool for fans and team management staff to predict the outcomes of NBA basketball games.

Keywords: NBA matches, Machine Learning, Pearson Correlation, Evaluation Metrics, Multivariate Regression, Linear Regression.

KATA PENGANTAR

Dengan mengucapkan rasa syukur kepada Tuhan Yang Maha Esa, penulis dapat menyelesaikan skripsi dengan judul "Prediksi Hasil Pertandingan Liga Bola Basket di NBA dengan Menggunakan Metode *Quadratic Multivariate*". Penelitian ini dilakukan dengan tujuan untuk memprediksi hasil pertandingan bola basket di NBA dengan metode yang akurat dan memberikan hasil yang lebih baik dibandingkan dengan metode lain. Penulis sangat berterima kasih kepada semua pihak yang telah membantu dalam penyelesaian skripsi ini, khususnya ucapan terimakasih ini kepada :

1. Prof. IGN Wiratmaja Puja Selaku Ketua Universitas Pertamina.
2. Budi W. Soetjipto, Ph.D. Selaku Wakil Ketua.
3. Muhamad Koyimatu, Ph.D Selaku Ketua Jurusan Ilmu Komputer.
4. Rangga Ganzar Noegraha, Ph.D Sebagai Dosen Pembimbing 1 Skripsi yang telah menyediakan waktu, energi dan pikirannya untuk memberikan bimbingan dan arahan selama proses penyusunan skripsi .
5. Ade Irawan, Ph.D selaku Dosen Pembimbing 2 Skripsi yang memberikan pengarahan dan bimbingan selama proses penyusunan skripsi.
6. Kedua Orang Tua saya yang telah lebih dulu meninggal duni ini dan selalu mendukung saya dalam hati dari awal saya menjelahi dunia pendidikan perkuliahan.
7. Kedua wali saya yang selalu mendukung dan membiayai saya untuk berkuliah.
8. Hakam Raves, Fernando, Fritz abytar dan Fauzi Arda yang selalu memberikan bantuan dan semangat dalam perjalanan skripsi.
9. Seluruh mahasiswa ilmu komputer Universitas Pertamina lainnya yang telah sama-sama memberikan bantuan dan semangat selama perkuliahan.
10. Dengan demikian penulis mengakhiri dengan mengucapkan terima kasih yang sebesar-besarnya. Terima kasih.

Serang, 13 Februari 2023



Boy Cakraningrat

DAFTAR ISI

KATA PENGANTAR	iii
DAFTAR ISI	iv
DAFTAR TABEL	vi
DAFTAR GAMBAR	vii
BAB I PENDAHULUAN	2
1.1 Latar Belakang	2
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA	5
2.1 Machine Learning	5
2.1.1 <i>Features Selection</i>	6
2.1.2 Model Regresi	7
2.2 Penelitian Terkait	9
BAB III METODE PENELITIAN	15
3.1 Pendekatan dan Metode Penelitian	15
3.2 Teknik Pengumpulan data	15
3.3 Pembuatan Model Prediksi Menggunakan <i>Machine Learning</i>	15
BAB IV HASIL DAN PEMBAHASAN	18
4.1 Pengumpulan Dataset	18
4.2 Pre-processing Data set	19
4.3 Seleksi Fitur	20
4.4 Membuat Model Prediksi dan Evaluasi	20

BAB V KESIMPULAN DAN SARAN	27
5.1 Kesimpulan	27
5.2 Saran	27
DAFTAR PUSTAKA	29
LAMPIRAN A Kumpulan data mentah nba season 2021-2022 dan code	42



1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah keceandekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.



DAFTAR TABEL

Tabel 2.1	Penelitian Sejenis	10
Tabel 4.1	Data Mentah Box Score Pertandingan NBA	18
Tabel 4.2	Hasil data set	19
Tabel 4.3	Hasil <i>Pearson</i> fitur seleksi	20
Tabel 4.4	Fitur pada dataset kolom 18 dengan korelasi tertinggi	20
Tabel 4.5	Fitur pada dataset kolom 22 dengan korelasi tertinggi	20
Tabel 4.6	Hasil evaluasi Uji model <i>Total point</i> dan <i>Opponent point</i> prediksi dataset 18 .	21
Tabel 4.7	Hasil evaluasi uji model <i>Total point</i> dan <i>Opponent point</i> prediksi dataset 22 .	21
Tabel 4.8	Hasil keakuratan model prediksi Dataset 18 dan Dataset 22	22
Tabel 4.9	Hasil bentuk dataset linear regresi	22
Tabel 4.10	Hasil evaluasi uji model <i>total point</i> dan <i>opponent point</i> prediksi dataset 18 . .	22
Tabel 4.11	Hasil evaluasi uji model <i>total point</i> dan <i>opponent point</i> prediksi dataset 22 . .	23
Tabel 4.12	Hasil keakuratan model prediksi Dataset 18 dan 22 metode regresi linear . .	23
Tabel 4.13	Perbandingan keakuratan antara model prediksi	23
Tabel 4.14	Perbedaan klasemen asli dan model prediksi musim lalu	24
Tabel 4.15	Akurasi 2 metode model	25



DAFTAR GAMBAR

Gambar 2.1	Flow <i>Machine Learning</i> (Chibani and Coudert, 2020)	5
Gambar 3.1	Diagram Alir Perancangan <i>Machine Learning</i>	15
Gambar 4.1	Hasil klasemen asli dan model prediksi <i>Quadratic Multivariate</i>	24
Gambar 4.2	Hasil prediksi Klasemen musim 2022/2023	25
Gambar A.1	Data kumpulan data	42

2. Dilarang mengutip karya tulis ini, kecuali:
 a. menyebutkan sumber sesuai kaidah keceandekiaan;
 b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 c. pengutipan tidak merugikan Universitas Pertamina.
 2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.



Universitas
Pertamina

Hak Cipta milik Universitas Pertamina

Dilindungi Undang-Undang

© Copyright of Universitas Pertamina



1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah kecependekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

Universitas
Pertamina



BAB I

PENDAHULUAN

1.1 Latar Belakang

National Basketball Association (NBA) adalah liga nasional bola basket di Amerika Serikat (AS) dengan pendapatan keuntungan terbesar di dunia. NBA memiliki pendapatan keuntungan yang lebih besar dari liga nasional bola basket lain, seperti *The Association of Basketball Clubs* (ACB) atau Liga Endesa Spain, dan Euro League atau Turkish Airlines. Pada musim 2019/2020, ACB memperoleh pendapatan keuntungan sebesar 30.27 juta Euro (2playbook, 2021) dan Euro League memiliki pendapatan 35 juta euro (Uğur Yılmaz, 2022). Sedangkan NBA sendiri memiliki pendapatan sekitar 7.9 miliar Dolar AS (Dimitrije Curcic, 2021).

Pada tahun 2021 terdapat 82 pertandingan *regular season* pada liga NBA. Setiap tim bola basket berusaha untuk memenangkan pertandingan *regular season* dan masuk *playoff* karena hal ini dapat meningkatkan pendapatan tim tersebut (Battista and Alojsio, 2019). Untuk mencapai hal tersebut, manajemen organisasi tim bola basket berusaha mengembangkan proses analisis kinerja olahraga yang efektif. *Machine Learning* merupakan alternatif solusi yang berfokus pada penggunaan data untuk proses analisa dan prediksi berdasarkan pola yang diperoleh dari data (Cholissodin and Imam, 2019). Dengan menggunakan jejak data tim pertandingan bola basket, *machine learning* digunakan untuk membantu manajemen, staf, pemain, dan dewan untuk membuat keputusan yang diperhitungkan secara maksimal dengan sedikit risiko.

Houston Rockets termasuk salah satu tim bola basket di liga NBA yang menggunakan jejak data pertandingan bola basket dengan hasil berupa keputusan untuk memaksimalkan poin per penguasaan bola dengan melakukan lebih banyak tembakan tiga poin dan hanya menembak di dalam kotak yang kemungkinan mendapatkan poinnya sangat tinggi. Pada musim 2014/2015, Houston Rockets menerapkan interpretasi analisis *Big Data*. Hasilnya, tim tersebut lolos menuju *play off* dengan peringkat pertama di *Southwest Division* dan peringkat kedua di *Western Conference*. Jika melihat statistik pertandingan tim pada musim 2014/2015, Houston Rockets lebih banyak menembak pada area tiga *point* dan dalam kotak tengah 2 *point*. Para pemainnya hampir tidak melakukan tembakan jarak menengah. Tren mengambil lebih banyak tembakan tiga angka ini selanjutnya menyebar ke seluruh tim di liga NBA (Corella Parra, 2021).

Sudah ada beberapa metode prediksi dalam pertandingan bola basket yang diusulkan para peneliti, seperti model Support Vector Machine (SVM), Random Forest, dan *polynomial regression* (Hsu et al., 2018). Belum ada yang mengusulkan penggunaan *Multivariate Quadratic Regression*. *Multivariate Quadratic Regression* adalah model yang digunakan untuk memprediksi banyak nilai variabel independen yang memiliki struktur non-linier (Sinha, 2013). Metode *multivariate* ini sudah pernah dipakai pada cabang olahraga lain, seperti pada penelitian (Young and Starkes, 2005) yang menyarankan bahwa latihan yang berkepanjangan sangat penting untuk pemeliharaan kinerja atletik bahkan dalam menghadapi prediksi penurunan terkait usia. Penelitian tersebut menggunakan analisis *Multivariate Quadratic Regression* untuk menguji hubungan antara usia dan performa lari pada nomor 1,500 dan 10,000 meter. Penelitian tersebut juga membandingkan usia dan penampilan sepanjang karir untuk 15 atlet pria asal Kanada.

Oleh karena itu, penelitian pada tugas akhir ini mengusulkan penggunaan *Multivariate Quadratic Regression* untuk prediksi poin tim dan poin tim lawan pada pertandingan bola basket. Model ini dibandingkan dengan model regresi *linear* dengan menggunakan data pertandingan bola basket yang sama, yaitu pada musim 2021/2022. Kedua model prediksi tersebut selanjutnya dievaluasi lalu modelnya digunakan untuk prediksi hasil pertandingan pada musim 2022/2023.

1.2 Rumusan Masalah

Dari sekian banyak usulan prediksi hasil akhir bola basket, belum ada yang menggunakan model *Multivariate Quadratic Regression* dan juga belum ada yang menunjukkan fitur pengaruh performa pemain terhadap prediksi hasil akhir tim bola basket. Sehingga, rumusan masalah untuk penelitian ini adalah :

1. Bagaimana hasil penerapan regresi *Multivariate Quadratic Regression* pada dalam melakukan prediksi performa pada setiap tim bola basket di NBA ?
2. Fitur apa saja yang sesuai untuk digunakan dalam memprediksi hasil pengaruh *stats* pemain terhadap total point tim dan point lawan tim.

1.3 Tujuan Penelitian

Berdasarkan uraian latar belakang, permasalahan yang harus diselesaikan dalam penelitian ini antara lain:

1. Mengetahui performa prediksi menggunakan metode regresi *Multivariate Quadratic Regression*, dengan hanya menggunakan data hasil pertandingan.
2. Mengetahui fitur hasil pengaruh *stats* pemain terhadap hasil akhir total poin tim dan poin tim lawan.

1.4 Batasan Masalah

Penelitian ini dilakukan untuk membangun model *machine learning* untuk memprediksi total *score* yang dicetak setiap tim dan total *score* yang dicetak lawannya dalam setiap pertandingan bola basket NBA pada musim 2022/2023 dengan menggunakan data hasil pertandingan di musim 2021/2022.

1.5 Manfaat Penelitian

1. Hasil penelitian ini dapat diduplikasi pada liga bola basket lainnya atau pada cabang olahraga lain.
2. Hasil penelitian ini bisa di manfaatkan oleh manajemen tim bola basket untuk menentukan tipe pemain yang lebih memungkinkan untuk memenangkan pertandingan.



Universitas
Pertamina

Hak Cipta milik Universitas Pertamina

Dilindungi Undang-Undang

© Copyright of Universitas Pertamina



1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah kecependekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

Universitas
Pertamina

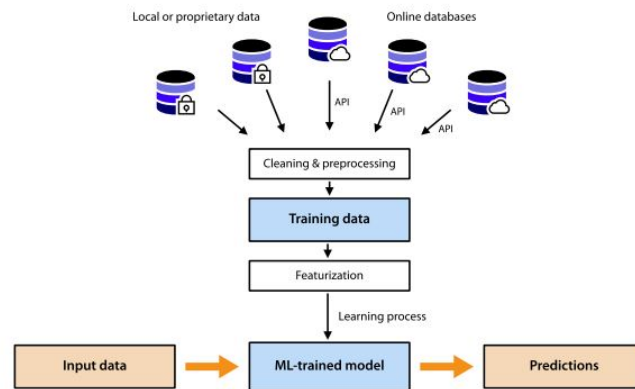


BAB II

TINJAUAN PUSTAKA

2.1 Machine Learning

Machine Learning adalah cabang dari pembelajaran *Artificial Intelligence* (AI) yang memungkinkan komputer untuk belajar dari data secara mandiri tanpa harus diprogram secara berulang-ulang. Dalam *Machine Learning*, algoritma dan proses statistika digunakan untuk melatih komputer dalam menemukan pola dan fitur tertentu dalam jumlah data yang besar, dengan tujuan untuk membuat keputusan atau prediksi berdasarkan data tersebut. Semakin baik bentuk algoritma, semakin tinggi pula akurasi keputusan dan prediksi yang dihasilkan oleh sistem (Cholissodin and Imam, 2019).



Gambar 2.1. Flow *Machine Learning* (Chibani and Coudert, 2020)

Dalam makalahnya di tahun 2020, Chibani and Coudert membagi tahapan pada *Machine Learning* sebagai berikut.

1. Mengumpulkan data (*Collectiong data*)

Proses mengumpulkan data atau *Gathering data* adalah langkah pertama dari alur kerja *Machine Learning* yang mencakup identifikasi, pengumpulan, atau pembuatan suatu set data atau dataset, dataset akan sangat bergantung pada tujuan model yang ingin dilatih.

2. *Cleaning and preprocessing data*

Proses membersihkan dan pra-proses kumpulan dataset besar mengidentifikasi kemungkinan titik data yang salah, hilang, atau tidak konsisten (*outlier*). Empat faktor penting dalam *cleaning* dan *preprocessing* data adalah sebagai berikut :

- Volume, yaitu jumlah dataset yang tersedia;
- Keberagaman, yaitu keanekaragaman atau heterogenitas data;
- Kejujuran, yaitu terkait pada pengetahuan tentang ketidakpastian data yang terkait dengan setiap titik dataset

(d) Kecepatan, yaitu seberapa cepat data dihasilkan dan harus diselesaikan.

Dari 4 faktor tersebut dapat disimpulkan bahwa data harus dihomogenkan dan dibersihkan sebelum dapat digunakan. Ini berarti mengidentifikasi kemungkinan titik dataset yang salah, hilang, atau tidak konsisten (Chibani and Coudert, 2020).

3. *Representing data*

Dataset yang dikumpulkan terkadang dalam bentuk format mentah dan perlu diubah menjadi format yang sesuai untuk tujuan model *Machine Learning* dalam rangkaian variabel skalar atau vektor untuk setiap entri kumpulan dataset. Langkah ini dilakukan untuk mengubah dataset sesuai tujuan model dengan transformasi data yang ada dengan penskalaan ulang data yang dibutuhkan. Normalisasi atau binarisasi bisa digunakan untuk menyesuaikan bentuk dataset sedemikian rupa sehingga algoritma yang dapat dibangun dengan mudah (Chibani and Coudert, 2020).

4. *Machine learning models*

Proses *training* dan *test* dilakukan untuk membuat modul algoritma *Machine Learning* dengan menggunakan kumpulan dataset yang telah dipreproses sebelumnya (Chibani and Coudert, 2020).

5. *Model evaluation*

Tujuan *Machine Learning* adalah melatih data dari dataset dan menghasilkan model komputasi yang efisien dan memiliki prediksi akurat. Keakuratan model diperiksa untuk mengevaluasi sejauh mana model menangkap dengan benar pola dasar dalam data, tetapi tidak dapat divalidasi hanya berdasarkan hasil yang diperoleh dari kumpulan data pelatihan tersebut (Chibani and Coudert, 2020).

2.1.1 *Features Selection*

Features Selection adalah metode yang digunakan dalam *Machine Learning* untuk mengidentifikasi fitur yang digunakan untuk mengurangi dimensi data, menyederhanakan model, meningkatkan interpretasi hasil model prediksi, dan meningkatkan performa model.

Pearson Selection

Koefisien Korelasi Pearson merupakan hubungan linear antara dua variabel acak yang memiliki nilai-nilai real. Koefisien korelasi linear ini digunakan untuk merefleksikan hubungan linear antara dua variabel kontinu yang normal (Jebli et al., 2021).

Nilai Korelasi Pearson dapat dihitung menggunakan persamaan berikut.

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

Di mana $\rho_{X,Y}$ adalah koefisien Korelasi Pearson, $\text{cov}(X, Y)$ adalah *covarians* antara dua variabel X dan Y , dan σ_X dan σ_Y adalah deviasi standar dari dua variabel. Koefisien ini memiliki range dari -1 sampai 1, di mana -1 menunjukkan korelasi negatif yang kuat, 1 menunjukkan korelasi positif yang kuat, dan 0 menunjukkan tidak ada korelasi antara dua variabel (Jebli et al., 2021).

2.1.2 Model Regresi

Regresi

Regresi adalah metode memperkirakan hubungan dari data yang diberikan untuk menggambarkan sifat kumpulan data. Hubungan ini kemudian dapat digunakan untuk berbagai perhitungan seperti prediksi nilai masa depan atau menghitung apakah ada hubungan antara berbagai variabel atau tidak (Sinha, 2013). Tahapan Model analisis regresi pada dasarnya terdiri dari empat macam tahapan:

- (a) Identifikasi variabel dependen dan independen.
- (b) Identifikasi bentuk hubungan antar variabel seperti linear, parabolik, eksponensial, dll dengan diagram pencar antara variabel dependen dan independen.
- (c) Perhitungan persamaan regresi untuk analisis.
- (d) Analisis kesalahan untuk memahami seberapa baik model estimasi sesuai dengan kumpulan data aktual (Sinha, 2013).

Berapa metode analisis regresi yang dikenal luas antara lain : Regresi linear, Regresi linear Multivariat, Regresi Polinomial, Regresi Polinomial Multivariat (Sinha, 2013).

Regresi Polinomial

Regresi Polinomial atau *Polynomial regression* adalah metode regresi *non-linear* dari data pengolahan data nilai variabel dependen yang meningkat atau menurun secara linear atau tampak parabola ketika data dikonstruksikan sebagai pencar (hubungan antara variabel dependen dan independen dikuadratkan). Persamaan dari Regresi Polinomial dapat diekspresikan sebagai :

$$a + b_1X + b_2X^2 + \dots b_pX^p \quad (2.1)$$

di mana konstanta a, b_1, b_2, \dots, b_p adalah parameter yang akan diestimasi. Contoh dari model regresi polinomial yang paling sederhana adalah model kuadrat.

$$Y = a + b_1X + b_2X^2 + e \quad (2.2)$$

Bentuk fungsi kuadrat $a + b_1X + b_2X^2$ adalah bentuk parabola (Severini, 2020).

Multivariate Polynomial

Regresi Polinomial dapat diterapkan pada regresi variabel prediktor tunggal yang disebut Regresi Polinomial Sederhana atau bisa juga dihitung pada regresi prediktor variabel berganda sebagai Regresi Polinomial Berganda (*Multivariate Polynomial Regression*).

Regresi multivariat polinomial berderajat 2 (*Multivariate Quadratic Regression*) dapat dinyatakan sebagai:



$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{11} x_{i1}^2 + \beta_{22} x_{i2}^2 + \beta_{12} x_{i1} x_{i2} + e_i \quad (2.3)$$

di mana :

- (a) $\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$ merupakan bentuk linear parameter efek
- (b) $\beta_{11} x_{i1}^2 + \beta_{22} x_{i2}^2$ merupakan bentuk parameter efek *quadratic*
- (c) $\beta_{12} x_{i1} x_{i2} + e_i$ merupakan bentuk *cross product* atau interaksi parameter efek. (Azubuiké, 2019).

Evaluasi Model Menggunakan R^2

R^2 adalah sebuah ukuran yang menunjukkan seberapa baik model memprediksi data yang ada. R^2 bernilai antara 0 dan 1, dengan nilai yang lebih tinggi menunjukkan bahwa model memprediksi data dengan lebih baik.

Formula untuk R^2 adalah sebagai berikut:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.4)$$

Di mana:

- (a) y_i adalah data aktual
- (b) \hat{y}_i adalah prediksi dari model
- (c) \bar{y} adalah nilai rata-rata dari data aktual n adalah jumlah data yang ada (Oytun et al., 2020).

Evaluasi Model Menggunakan *Mean Absolute Error (MAE)*

Mean Absolute Error (MAE) adalah ukuran error rata-rata yang mengukur seberapa besar kesalahan prediksi yang dilakukan oleh model. MAE dalam rumus matematis dapat dituliskan sebagai berikut:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.5)$$

Di mana:

- (a) y_i adalah data aktual
- (b) \hat{y}_i adalah prediksi dari model
- (c) n adalah jumlah data yang ada
- (d) $|y_i - \hat{y}_i|$ adalah *absolute error* antara data aktual dan prediksi (Oytun et al., 2020).



2.2 Penelitian Terkait

Beberapa penelitian yang terkait dengan penelitian ini adalah sebagai berikut :

- (a) Studi (Young and Starkes, 2005) menyarankan bahwa latihan yang berkepanjangan sangat penting untuk pemeliharaan performa atletik bahkan dalam menghadapi prediksi penurunan terkait usia. Penelitian ini menggunakan analisis regresi polinomial berganda *multivariate polynomial* untuk menguji hubungan antara usia dan performa lari pada nomor 1500 dan 10.000 meter. Penelitian ini membandingkan usia dengan penampilan sepanjang karir untuk 15 atlet pria asal Kanada.
- (b) Penelitian (Imran et al., 2022) bertujuan untuk membuat model prediksi kuat tekan campuran bahan beton. Pada penelitian ini dibangun model prediksi campuran beton yang mengandung beton agregat daur ulang (RAC) dan terak tanur tiup butiran tanah (GGBFS). Model *Machine Learning* menggunakan model regresi polinomial multivariat (MPR) digunakan untuk memprediksi kekuatan tekan beton ramah lingkungan. Model tersebut dibandingkan dengan dua model *Machine Learning* lainnya, yaitu regresi linear (LR) dan model *support vector machine* (SVM). 3 Model tersebut di uji dengan menggunakan R^2 (koefisien determinasi) dan RMSE (root mean absolute error).
- (c) Penelitian (Stathopoulos and Karlaftis, 2003) bertujuan untuk mengembangkan model ruang keadaan waktu multivariat yang fleksibel, dengan menggunakan *score urban area loop detector* dan data detektor lingkaran kota utama. Data yang digunakan dalam penelitian ini adalah data jalan arteri perkotaan yang berdekatan dengan pusat kota Athena, untuk memprediksi model deret waktu multivariat secara eksplisit. Hasil penelitian menunjukkan bahwa spesifikasi yang berbeda cocok untuk periode waktu yang berbeda di Athena, dan penggunaan model *time series multivariate* memiliki potensi dalam sistem jalan raya perkotaan.
- (d) Penelitian (Hsu et al., 2018) bertujuan untuk membuat model prediksi liga bola basket di NBA dengan beberapa pendekatan model yang berbeda seperti model regresi *support vector machine*, model regresi polinomial, dan model regresi *Random forest regression* sebagai analisis komparatif untuk menemukan bagaimana kinerja pemain individu memengaruhi tingkat kemenangan tim.

Untuk keempat penelitian tersebut pada Tabel 2.1 dilakukan analisis dengan menggunakan 3C2S (*Compare, Contrast, Criticize, Synthesize, Summarize*)

Dilarang mengutipnya tulis ini, kecuali:

- a. menyebutkan sumber sesuai kaidah kecedekiaan;
- b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
- c. pengutipan tidak merugikan Universitas Pertamina.

2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

Tabel 2.1. Penelitian Sejenis

Jurnal	<i>Compare</i>	<i>Contrast</i>	<i>Criticize</i>	<i>Synthesize</i>	<i>Summarize</i>
Jurnal 1: <i>Career-span analyses of track performance: longitudinal data present a more optimistic view of age-related performance decline. Experimental Aging Research</i>	Penelitian ini melakukan prediksi dengan menggunakan model regresi polinomial multivariat (MPR) <i>quadratic</i> .	Penelitian ini menggunakan tiga model analisis regresi untuk masing-masing set data cross-sectional dan longitudinal: model linear, model eksponensial, dan model polinomial orde kedua.	Masing-masing model diperiksa dengan menggunakan R^2 untuk mengevaluasi kinerja model. Nilai R^2 berbeda antara penerapannya pada data <i>cross-sectional</i> , dan pada data karir <i>longitudinal</i> .	Penggunaan 3 Model regresi menunjukkan bahwa performa lari pada karir <i>longitudinal</i> menurun seiring bertambahnya usia. Penurunan ini memiliki pola yang lebih linear dibandingkan data <i>cross-sectional</i> .	Penggunaan 3 model regresi dataset <i>cross-sectional</i> dan data longitudinal karir menunjukkan bahwa rentan pelatihan karir berperan dalam performa kinerja lari atlet elit.

Jurnal	<i>Compare</i>	<i>Contrast</i>	<i>Criticize</i>	<i>Synthesize</i>	<i>Summarize</i>
Jurnal 2: <i>Development of prediction model to predict the compressive strength of eco-friendly concrete using multivariate polynomial regression combined with step-wise method</i>	Penelitian ini melakukan prediksi dengan menggunakan model regresi polinomial multivariat <i>polynomial multivariat quadratic</i> .	Penelitian ini menggunakan model regresi polinomial multivariat untuk memprediksi kekuatan tekan beton ramah lingkungan. Dibandingkan dengan dua model lainnya yaitu regresi linear, dan model <i>support vector machine</i> . 3 Model tersebut dievaluasi menggunakan pengukuran R^2 dan RMSE.	Penelitian ini memiliki keterbatasan dataset yang kecil digunakan untuk membuat model prediksi, hanya menggunakan 125 kasus hal ini membuat model prediksi <i>Machine Learning</i> kurang bagus.	Dalam penelitian ini model <i>Machine Learning</i> digunakan untuk memprediksi kekuatan tekan beton ramah lingkungan yang mengandung RAC dan GGBFS. Model MPR menunjukkan performa yang lebih baik dibandingkan dengan model LR dan SVM dalam pengujian. Namun, evaluasi menggunakan R^2 dan RMSE memiliki kelemahan, sehingga perlu diper-timbangkan evaluasi lain.	Pada penelitian ini ketiga model regresi, yaitu model regresi polinomial multivariat, regresi linear, dan model <i>Support Vector Machine</i> (SVM), digunakan untuk memprediksi kekuatan tekan beton ramah lingkungan. Setelah diuji dengan menggunakan metrik evaluasi R^2 dan RMSE, ditemukan bahwa regresi linear memiliki performa yang paling buruk dibandingkan dengan kedua model lainnya.

Jurnal	<i>Compare</i>	<i>Contrast</i>	<i>Criticize</i>	<i>Synthesize</i>	<i>Summarize</i>
Journal 3: A <i>multivariate state space approach for urban traffic flow modeling and prediction</i> .	Penelitian ini membangun model <i>multivariate time series</i> bervariasi yang fleksibel dengan menggunakan <i>core urban area loop detector</i> (data detektor lingkaran kota utama).	Pembuatan model <i>multivariate time series</i> dilakukan dengan mengukur volume jalan arteri selama 3 menit di dekat pusat kota Athena. model ini dibangun dengan memperhitungkan data dari detektor hulu untuk memperbaiki prediksi lokasi hilir.	Penelitian ini bertujuan untuk memodelkan kondisi lalu lintas padat dan tidak stabil. Namun, hasil dari pemodelan statistika yang dilakukan kurang akurat dalam memprediksi karakteristik lalu lintas jangka pendek. Oleh karena itu, diperlukan pendekatan baru yang dapat mengakomodasi perilaku lalu lintas pada kondisi terbatas.	Hasil penelitian ini menunjukkan pembuatan model prediksi <i>multivariate time series</i> yang fleksibel untuk lalu lintas perkotaan kota Athena dengan menggunakan data detektor lingkaran kota utama. Namun, hasil penelitian menunjukkan adanya keterbatasan dalam generalisasi temuan wilayah lalu lintas perkotaan. Oleh karena itu, perlu dilakukan penelitian lebih lanjut untuk memperluas generalisasi kondisi lalu lintas perkotaan lainnya dan menguji model dalam situasi lalu lintas yang lebih kompleks.	Hasil dari model yang dikembangkan menunjukkan jelas bahwa, setidaknya dalam kasus Athena, spesifikasi yang berbeda sesuai untuk periode waktu yang berbeda. Selain itu, terlihat bahwa penggunaan model <i>time series multivariate</i> memberikan potensi sistem alat yang efektif dalam mengelola lalu lintas perkotaan dalam sistem jalan raya perkotaan.

1. Dilarang mengutip karya tulis ini, kecuali:

- menyebutkan sumber sesuai kaidah kependidikan;
- pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;

c. pengutipan tidak merugikan Universitas Pertamina.

2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

Jurnal	<i>Compare</i>	<i>Contrast</i>	<i>Criticize</i>	<i>Synthesize</i>	<i>Summarize</i>
Journal 4: <i>Evaluating Machine Learning Techniques for NBA Players' Winning Contribution.</i>	Memiliki persamaan kasus melakukan prediksi menggunakan pendekatan model yang berbeda, yaitu model regresi <i>support vector machine</i> , model regresi polinomial, dan model regresi <i>Random forest Regression</i> , untuk menganalisis bagaimana kinerja pemain individu memengaruhi tingkat kemenangan tim di NBA	Model prediksi yang digunakan adalah regresi <i>Support Vector</i> , regresi polinomial, dan regresi hutan acak (<i>Random forest regression</i>). Penelitian ini melakukan analisis komparatif untuk tiap model tersebut	Hasil dari model penelitian prediksi ini tidak diuji dengan data musim selanjutnya.	Penelitian ini melakukan prediksi dengan menggunakan 3 model regresi, yaitu <i>support vector Machine Learning</i> , regresi polinomial, dan <i>Random forest Regression</i> untuk mengetahui model paling baik yang digunakan untuk prediksi kinerja tim.	Pada penelitian ini pembuatan model prediksi tim bola basket NBA menggunakan data musim tahun 2017-2018 dengan metode regresi <i>support vector Machine Learning</i> , regresi polinomial, dan <i>Random forest regression</i> .



Universitas
Pertamina

Hak Cipta milik Universitas Pertamina

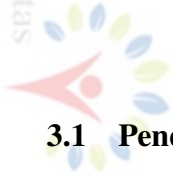
Dilindungi Undang-Undang

© Copyright of Universitas Pertamina



1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah kecependekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

Universitas
Pertamina



BAB III

METODE PENELITIAN

3.1 Pendekatan dan Metode Penelitian

Penelitian ini dilakukan dengan menggunakan metode penelitian kuantitatif untuk mengumpulkan dan menganalisis informasi atau data dalam bentuk angka. Tujuan dari penggunaan metode kuantitatif adalah untuk menguji hasil dari model hipotesis yang telah ditetapkan. Alat penelitian yang digunakan dalam penelitian ini adalah teknik statistik menggunakan *Machine Learning* untuk menganalisis data secara kuantitatif.

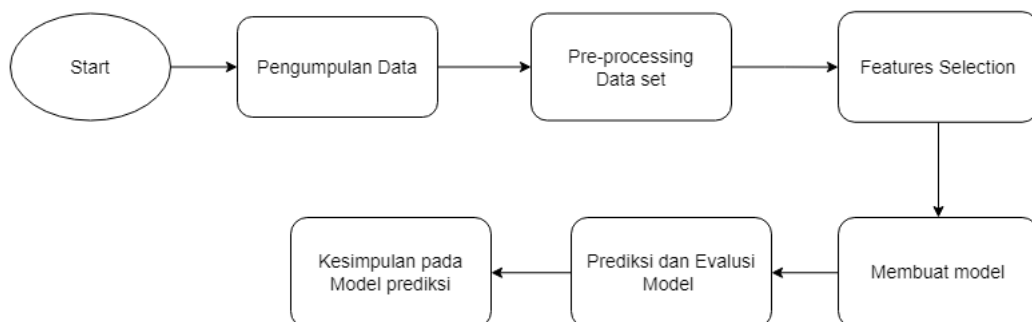
3.2 Teknik Pengumpulan data

Pengumpulan Data yang dilakukan menggunakan teknik observasi untuk menentukan spesifikasi kebutuhan model prediksi *Machine Learning*. Spesifikasi kebutuhan data berisi poin-poin yang harus tersedia pada data yang akan dikumpulkan untuk membuat model prediksi (Id, 2021). Pengumpulan data pertandingan bola basket menggunakan teknik observasi dengan mencari sumber yang valid dan dapat dipertanggungjawabkan. Sumber data pertandingan bola basket diambil dari website *Basketball Sport Reference* yaitu website yang menyediakan data pemain, tim dan laga pertandingan bola basket di NBA.

3.3 Pembuatan Model Prediksi Menggunakan *Machine Learning*

Langkah yang dilakukan pada penelitian ini untuk membuat model prediksi ditampilkan pada diagram alir pada Gambar 3.1.

Gambar 3.1. Diagram Alir Perancangan *Machine Learning*



Pada langkah pertama yaitu pengumpulan data, yang harus dilakukan adalah mengumpulkan data mentah hasil pertandingan bola basket selama satu musim. Data mentah pertandingan didapatkan dari sumber yang valid sebagai bahan utama dalam pembuatan model prediksi.

Setelah data mentah dikumpulkan, langkah selanjutnya adalah melakukan preprocessing pada data. Data mentah dipreproses sehingga memiliki format yang tepat dan siap digunakan untuk membangun model. Setelah dataset dibuat, tahap selanjutnya adalah pemilihan fitur yang akan digunakan untuk membangun model. Pemilihan fitur dilakukan berdasarkan analisis korelasi Pearson.

Analisis korelasi Pearson dipilih sebagai metode untuk pemilihan fitur karena teknik ini dapat mengidentifikasi variabel yang memiliki hubungan kuat dengan target variabel yang akan diprediksi. Model yang dibangun pada penelitian ini, merupakan model regresi yang dibangun melalui pendekatan regresi linear. Sehingga nilai Korelasi Pearson menjadi nilai Korelasi yang sesuai untuk melakukan pemilihan fitur. Dengan demikian, variabel yang memiliki korelasi rendah atau tidak signifikan dengan target variabel dapat dihilangkan dari model untuk meningkatkan akurasi linear dan efisiensi model.

Setelah dilakukan pemilihan fitur pada model dataset, langkah selanjutnya dalam membuat model prediksi adalah menggunakan metode regresi *Quadratic Multivariate* dan metode pembandingan *Linear Regression*. Pembuatan model prediksi pada dataset dilakukan dengan memisahkan dataset menjadi data train sebesar 80 persen dan data test sebesar 20 persen . Data train digunakan untuk melatih model prediksi, sedangkan data test digunakan untuk mengevaluasi kinerja model yang telah dilatih (Nguyen et al., 2022).

Pemisahan data train dan data test dilakukan untuk memastikan bahwa model prediksi yang dibangun tidak hanya mampu melakukan prediksi pada data yang telah dikenal, tetapi juga mampu melakukan prediksi dengan akurat pada data yang belum dikenal. Dalam proses evaluasi model, kinerja model dievaluasi menggunakan beberapa metrik evaluasi, misalnya Root Mean Squared Error (RMSE), dan R-squared (R^2). Dengan melakukan evaluasi kinerja model menggunakan metrik-metrik tersebut, dapat diketahui apakah model yang dibangun sudah cukup baik atau masih perlu diperbaiki.

Penelitian ini membangun dua model, yaitu model dengan target *total point* dan model dengan target *opponent point*. Evaluasi keakuratan 2 model prediksi kemenangan pertandingan tim bola basket di NBA dihitung dengan menggunakan model prediksi *total point* dan *opponent point*. Model prediksi ini dibangun terdiri dari 3 jenis model prediksi. Dari ketiga model dicari keakuratan prediksi dengan membandingkan hasil pertandingan kemenangan tim bola basket dengan hasil kemenangan prediksi memenangkan pertandingan pada musim tahun 2021/2022 serta 2022/2023. Evaluasi performa model prediksi dilakukan dengan menghitung keakuratan nilai persentase perbandingan kesesuaian nilai prediksi dengan hasil pertandingan sebenarnya di musim 2021/2022. Model dengan nilai persentase prediksi terbesar digunakan untuk melakukan prediksi hasil pertandingan musim 2022/2023.



Universitas
Pertamina

Hak Cipta milik Universitas Pertamina

Dilindungi Undang-Undang

© Copyright of Universitas Pertamina



1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah kecependekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

Universitas
Pertamina

BAB IV

HASIL DAN PEMBAHASAN

4.1 Pengumpulan Dataset

Data yang digunakan untuk memprediksi pertandingan bola basket adalah data pertandingan tim yang diambil pada website <https://www.basketball-reference.com> sebanyak 1323 data pertandingan di musim 2021-2022. Data yang diambil dari website tersebut berupa hasil pertandingan (*box score*) pertandingan NBA antara 2 tim dalam 1 pertandingan. Dataset telah disimpan pada url https://github.com/boycakra/NBA_Prediction isi fitur pengumpulan data mentah diperlihatkan pada.

Tabel 4.1. Data Mentah Box Score Pertandingan NBA

Nama Fitur	Keterangan
Starters	Berisikan nama pemain tim bola basket yang bermain awal.
Mp	<i>Minute played</i> berisikan waktu pemain bermain pada pertandingan tersebut.
Fg	<i>Feild goal</i> berisikan total <i>score</i> yang dicetak pemain selain <i>free throw</i> .
Fga	<i>Feild goal attempt</i> berisikan banyaknya percobaan pemain untuk mencetak angka dalam pertandingan tersebut.
Fg%	<i>Feild goal percentage</i> berisikan persentase berhasilnya pemain dalam mencetak angka dalam pertandingan tersebut.
Ft	<i>Free throw</i> berisikan total <i>score</i> yang dicetak pemain dari tembakan bebas.
Fta	<i>Free throw attempt</i> berisikan banyaknya percobaan tembakan bebas.
Ft%	<i>Free throw %</i> berisikan persentase tembakan bebas yang berhasil
Orb	<i>Offensive rebound</i> berisikan data seorang pemain melakukan pengambilan pantulan bola basket pada tiang ring musuh.
Drb	<i>Defensive Rebound</i> kolom ini berisikan data seorang pemain melakukan pengambilan pantulan bola basket pada tiang ring team.
Trb	<i>Total Defensive Rebound</i> berisikan jumlah Orb dan Drb.
Ast	<i>Assist</i> berisikan data seorang pemain melakukan bantuan terhadap rekan timnya untuk mencetak poin.
Steal	<i>Steal</i> berisikan data seorang pemain mencuri bola basket dari pemain musuh
Blk	<i>Block</i> berisikan data seorang pemain menangkal bola basket dari pemain musuh saat akan mencetak <i>score</i> .
Tov	<i>Turnover</i> berisikan data seorang pemain kehilangan bola dan diambil oleh pemain lawan.
Pf	<i>Personal foul</i> berisikan data seorang pemain melakukan pelanggaran dalam pertandingan.
Pts	<i>Points</i> berisikan data total <i>score</i> yang dicetak pemain dalam pertandingan.
Plus/minus	Berisikan data perbedaan skor ketika pemain masuk lapangan sampai dengan meninggalkan lapangan.

4.2 Pre-processing Data set

Dari kumpulan data mentah yang telah dikumpulkan dilakukan *pre-process* data dengan tujuan mengubah format data sesuai dengan yang dibutuhkan untuk membangun model. Proses *pre-processing* data secara terurut adalah sebagai berikut :

1. Mengganti nilai data kosong atau *null* karena data yang dibutuhkan adalah informasi pemain yang bermain pada pertandingan tersebut.
2. Mendrop kolom pada dataset yang tidak digunakan karena data yang dibutuhkan adalah informasi pengaruh pemain dalam pertandingan.
3. Menambahkan kolom tim, lawan tim beserta jumlah *score* karena data yang dibutuhkan adalah informasi mengenai pertandingan tersebut.
4. Menambahkan kolom *home/visitor* Untuk menunjukkan pengaruh perbedaan dalam data pada pertandingan ketika dimainkan di daerah sendiri atau daerah lawan.
5. Mengubah nilai kolom *mp (minute played)* menjadi waktu detik bermain untuk mengetahui pengaruh pemain dalam pertandingan dengan lebih akurat.
6. Mengubah nilai kolom tim menjadi peringkat tahun lalu tim NBA untuk mengetahui pengaruh peringkat tim pada pertandingan.
7. Membagi data menjadi 2 dataset, Dataset 18 dan 22. Untuk mencari tahu pengaruh dari perbedaan model prediksi antara 2 dataset tersebut

Dari hasil *pre-processing* ini dihasilkan dua dataset isi pada tiap dataset kolom kolom percobaan, seperti pada Tabel 4.2.

Tabel 4.2. Hasil data set

Dataset	kolom
Dataset 18 kolom	Tanggal, Team, Home/Visitor, Starters, Position, MP, FGA, 3PA, FTA, ORB, DRB, STL, BLK, TOV, PF, Total Point, Opponent, Opponent Pts
Dataset 22 kolom	Tanggal, Team, Home/Visitor, Starters, Position, MP, FG, FGA, 3P, 3PA, FT, FTA, ORB, DRB, AST, STL, BLK, TOV, PF, Total Point, Opponent, Opponent Pts

Semua fitur pada Dataset 22 termuat pada Dataset 18, kecuali fitur *fg*, *3p*, *ft* dan *ast*. *pre-process* dilanjutkan lagi dengan mengambil data 5 pemain pada setiap tim bola basket *NBA* berdasarkan atas nilai waktu bermain paling besar di setiap pertandingan. Data 5 pemain ini diurutkan sesuai posisi. Proses selanjutnya adalah :

1. Mengurutkan dan mengambil pemain dengan nilai waktu bermain paling banyak, lalu diurutkan sesuai posisi pemain. Data set yang diinginkan adalah berisikan 5 pemain dengan jumlah waktu bermain paling banyak dalam satu pertandingan.

2. Mengabungkan 5 baris pemain menjadi 1 baris sesuai tim dan tanggal bermain. Data set yang diinginkan berisikan 1 baris yang memuat data statistik 5 pemain dalam satu pertandingan.
3. Menghilangkan kolom starter pada kedua dataset. Karena kolom starter tidak digunakan dalam pembuatan model prediksi.
4. Mengalikan setiap kolom dan membuat nama kolom baru.

4.3 Seleksi Fitur

Setelah melakukan *pre-processing* pada 2 dataset tersebut dilakukan seleksi fitur, seleksi fitur dilakukan dengan melihat hasil Korelasi *Pearson*, yang mengukur korelasi linier antara tiap fitur dataset dengan target yaitu kolom *total poin* dan *opponent point*. Hasil sebelum dan sesudah seleksi fitur diperlihatkan pada Tabel 4.3.

Tabel 4.3. Hasil *Pearson* fitur seleksi

Dataset	sebelum	<i>total poin</i>	<i>opponent poin</i>
Dataset 18 kolom	1710 kolom	50 kolom	74 kolom
Dataset 22 kolom	3080 kolom	793 kolom	93 kolom

Fitur-fitur pada 2 dataset yang memiliki nilai Korelasi *Pearson* tertinggi ditampilkan Tabel 4.4 dan Tabel 4.5.

Tabel 4.4. Fitur pada dataset kolom 18 dengan korelasi tertinggi

Total Points	nilai korelasi	Opponent Points	nilai korelasi
FGA1_FTA4	0.13234	Team_FGA	0.14001
Opponent_FGA5	0.12585	Team_FTA1	0.13066
FTA1_FTA2	0.12365	Team_3PA1	0.11739
FTA1_FTA4	0.12339	Team_FGA5	0.11657
FGA1_3PA4	0.12333	Team_PF4	0.11354

Tabel 4.5. Fitur pada dataset kolom 22 dengan korelasi tertinggi

Total Points	nilai korelasi	Opponent Points	nilai korelasi
FG1_FG5	0.30844	Team_FGA1	0.14001
AST4_AST5	0.30193	Team_FG1	0.13846
FG1_FG4	0.30189	Team_FG5	0.13478
FG5_AST5	0.29849	Team_FTA1	0.13066
FG1_AST5	0.29827	FG1_FG5	0.12881

4.4 Membuat Model Prediksi dan Evaluasi

Setelah melakukan seleksi fitur selanjutnya dilakukan pembuatan model. Dalam pembuatan model, dataset dibagi menjadi *train* data dan *test* data dengan perbandingan 80:20. *Train* data digunakan untuk melatih model prediksi sedangkan *test* data digunakan untuk mengevaluasi model (Nguyen et al., 2022).

Untuk setiap dataset dibangun 2 model prediksi, model prediksi *total poin* dan model prediksi untuk *opponent poin*. Hasil Model tersebut dievaluasi dengan menggunakan evaluasi metrik R^2 dan

MAE (*Mean Absolute Error*), kedua model tersebut diuji dengan data test dan dengan data train. Hasilnya dari evaluasi model ditampilkan pada Tabel 4.6 Dataset 18 dan untuk Dataset 22 pada Tabel 4.7

Tabel 4.6. Hasil evaluasi Uji model *Total point* dan *Opponent point* prediksi dataset 18

Evaluasi <i>Total point</i>	Hasil	Evaluasi <i>Opponent point</i>	Hasil
Hasil R^2 test	0.08209	Hasil R^2 test	0.21781
Hasil R^2 train	0.12107	Hasil R^2 train	0.22814
Hasil MEA	8.92800	Hasil Mea	9.07006

Tabel 4.7. Hasil evaluasi uji model *Total point* dan *Opponent point* prediksi dataset 22

Evaluasi <i>Total point</i>	Hasil	Evaluasi <i>Opponent point</i>	Hasil
Hasil R^2 test	0.41377	Hasil R^2 test	0.23121
Hasil R^2 train	0.80252	Hasil R^2 train	0.25712
Hasil Mea	18.11669	Hasil Mea	8.92423

Dari Tabel 4.6 dan Tabel 4.7 evaluasi metrik model prediksi *total point* 22 memiliki nilai evaluasi terbaik. Setelah melakukan test evaluasi metrik, selanjutnya dilakukan evaluasi prediksi dengan data sebenarnya. Data ini adalah data pertandingan di musim 2021-2022 yang memiliki 1323 pertandingan.

Dalam penelitian ini, terdapat 2 model tambahan yang dibangun dengan menggunakan kedua dataset tanpa melakukan split data, sehingga 100% data digunakan untuk training model. Model 3 dibangun dengan menggunakan dataset 18, sedangkan Model 4 dibangun menggunakan dataset 22. Total model yang dibangun pada penelitian ini adalah 6 model, dengan Model 5 sebagai modifikasi dari Model 1. Pada Model 5, hanya model prediksi Total Poin dari Model 1 yang digunakan. Nilai prediksi opponent point diambil dari nilai prediksi Total Poin tim lawan, dengan menggunakan hasil prediksi Total Poin Tim B saat melawan Tim A sebagai opponent point pada pertandingan Tim A melawan Tim B. Model 6 juga dibangun dengan cara yang sama, tetapi menggunakan Model 2. Ketepatan keenam model tersebut ditampilkan pada Tabel 4.8

Tabel 4.8. Hasil keakuratan model prediksi Dataset 18 dan Dataset 22

Nama	Split & test	Dataset	Model Prediksi Digunakan	Persentase Keakuratan
Model 1	80-20	18	Tp dan Op	0.67735
Model 2	80-20	22	Tp dan Op	0.61886
Model 3	100-0	18	Tp dan Op	0.66931
Model 4	100-0	22	Tp dan Op	0.75736
Model 5	80-20	18	Tp	0.60770
Model 6	80-20	22	Tp	0.84278

Hasil penelitian menunjukkan bahwa dengan menggunakan kedua dataset tanpa melakukan split data, akurasi model dapat ditingkatkan. Selain itu, Model 6 berhasil memperbaiki akurasi prediksi dengan hanya menggunakan model prediksi Total Poin dari Model 2. Dalam penelitian ini, penggunaan Model 6 memberikan alternatif penggunaan model yang efektif dalam memprediksi hasil pertandingan tim bola basket di NBA.

Sebagai perbandingan, pada penelitian ini dilakukan 6 model perbandingan menggunakan metode regresi linear dengan menggunakan semua fitur pada Tabel 4.9. Hasil evaluasi metrik linear untuk Dataset 18 dan Dataset 22 Tabel 4.10 dan Tabel 4.11.

Tabel 4.9. Hasil bentuk dataset linear regresi

Dataset	Fitur - fitur
Dataset 18 kolom	Team, Opponent, Home/Visitor1, MP1, FGA1, 3PA1, FTA1, ORB1, DRB1, STL1, BLK1, TOV1, PF1, Home/Visitor2, MP2, FGA2, 3PA2, FTA2, ORB2, DRB2, STL2, BLK2, TOV2, PF2, Home/Visitor3, MP3, FGA3, 3PA3, FTA3, ORB3, DRB3, STL3, BLK3, TOV3, PF3, Home/Visitor4, MP4, FGA4, 3PA4, FTA4, ORB4, DRB4, STL4, BLK4, TOV4, PF4, Home/Visitor5, MP5, FGA5, 3PA5, FTA5, ORB5, DRB5, STL5, BLK5, TOV5, PF5
Dataset 22 kolom	Team, Opponent, Home/Visitor1, MP1, FG1, FGA1, 3P1, 3PA1, FT1, FTA1, ORB1, DRB1, AST1, STL1, BLK1, TOV1, PF1, Home/Visitor2, MP2, FG2, FGA2, 3P2, 3PA2, FT2, FTA2, ORB2, DRB2, AST2, STL2, BLK2, TOV2, PF2, Home/Visitor3, MP3, FG3, FGA3, 3P3, 3PA3, FT3, FTA3, ORB3, DRB3, AST3, STL3, BLK3, TOV3, PF3, Home/Visitor4, MP4, FG4, FGA4, 3P4, 3PA4, FT4, FTA4, ORB4, DRB4, AST4, STL4, BLK4, TOV4, PF4, Home/Visitor5, MP5, FG5, FGA5, 3P5, 3PA5, FT5, FTA5, ORB5, DRB5, AST5, STL5, BLK5, TOV5, PF5

Tabel 4.10. Hasil evaluasi uji model *total point* dan *opponent point* prediksi dataset 18

Evaluasi <i>Total point</i>	Hasil	Evaluasi <i>Opponent point</i>	Hasil
Hasil R^2 test	0.10588	Hasil R^2 test	0.25542
Hasil R^2 train	0.15255	Hasil R^2 train	0.26095
Hasil Mea	8.83052	Hasil Mea	8.94650

Dari Tabel 4.10 dan 4.11 evaluasi matrix *Regresi linear* model total point 22 memiliki kecocokan terbaik dengan data, dengan nilai R^2 tertinggi train, test dan nilai MAE terendah. Selanjutnya

Tabel 4.11. Hasil evaluasi uji model *total point* dan *opponent point* prediksi dataset 22

Evaluasi <i>Total point</i>	Hasil	Evaluasi <i>Opponent point</i>	Hasil
Hasil R^2 test	0.61544	Hasil R^2 test	0.26315
Hasil R^2 train	0.68916	Hasil R^2 train	0.28749
Hasil Mea	5.77102	Hasil Mea	8.85394

dilakukan evaluasi prediksi kemenangan pertandingan pada setiap tim dengan data sebenarnya menggunakan model prediksi *total point* dan *Opponent point*.

Selanjutnya pada penelitian ini dibangun 6 model pembanding menggunakan regresi linear. 6 model pembanding tersebut ditampilkan pada Tabel 4.12.

Tabel 4.12. Hasil keakuratan model prediksi Dataset 18 dan 22 metode regresi linear

Pembanding	Split & test	Dataset	Model Prediksi Digunakan	Persentase Keakuratan
Model 1	80-20	18	Tp dan Op	0.72264
Model 2	80-20	22	Tp dan Op	0.76792
Model 3	100-0	22	Tp dan Op	0.68896
Model 4	100-0	22	Tp dan Op	0.77853
Model 5	80-20	18	Tp	0.62811
Model 6	80-20	22	Tp	0.83824

Dari Tabel 4.8 dan Tabel 4.12 ditemukan bahwa ketelitian paling tinggi untuk masing masing metode regresi adalah model pada Tabel 4.13

Tabel 4.13. Perbandingan keakuratan antara model prediksi

Model	Metode	Hasil persen model prediksi
Model 6	<i>Quadratic Multivariate</i>	0.84278
Model 6 pembanding	Regresi linear <i>Regresi linear</i>	0.83824

Dari hasil Tabel 4.13 dapat dilihat bahwa hasil dari metode *Quadratic Multivariate* untuk Model 6 memiliki hasil lebih bagus dari pada metode *Regresi linear*. Model 6 *Quadratic Multivariate* digunakan untuk memprediksi klasemen. Hasil prediksi klasemen dievaluasi dengan klasemen asli musim 2021/2021 yang digunakan dalam membuat model prediksi. Hasil dari prediksi klasemen menggunakan model 5 dan 6 pada tiap metode *Quadratic Multivariate* dan metode regresi linier musim 2021/2022 terdapat pada Gambar 4.1.

Gambar 4.1. Hasil klasemen asli dan model prediksi *Quadratic Multivariate*

R.ASLI	P.TP18bagi2	P.TP22bagi2	P.Linear Tp18_bagi 2	P. linear TP22_bagi 2
Phoenix Suns	Dallas Mavericks	Phoenix Suns	Philadelphia 76ers	Golden State Warriors
Memphis Grizzlies	Utah Jazz	Boston Celtics	Phoenix Suns	Phoenix Suns
Golden State Warriors	Milwaukee Bucks	Golden State Warriors	Utah Jazz	Dallas Mavericks
Miami Heat	Philadelphia 76ers	Memphis Grizzlies	Milwaukee Bucks	Philadelphia 76ers
Dallas Mavericks	Brooklyn Nets	Milwaukee Bucks	Memphis Grizzlies	Milwaukee Bucks
Boston Celtics	Atlanta Hawks	Dallas Mavericks	Dallas Mavericks	Boston Celtics
Milwaukee Bucks	Los Angeles Lakers	Philadelphia 76ers	Brooklyn Nets	Miami Heat
Philadelphia 76ers	Memphis Grizzlies	Miami Heat	Atlanta Hawks	Memphis Grizzlies
Utah Jazz	Los Angeles Clippers	Toronto Raptors	Golden State Warriors	Denver Nuggets
Toronto Raptors	Golden State Warriors	Utah Jazz	Los Angeles Lakers	Brooklyn Nets
Denver Nuggets	Phoenix Suns	Denver Nuggets	Toronto Raptors	Chicago Bulls
Chicago Bulls	Portland Trail Blazers	Chicago Bulls	Portland Trail Blazers	Utah Jazz
Minnesota Timberwolves	Toronto Raptors	Los Angeles Clippers	Denver Nuggets	San Antonio Spurs
Cleveland Cavaliers	New York Knicks	Minnesota Timberwolves	Minnesota Timberwolves	Charlotte Hornets
Brooklyn Nets	Denver Nuggets	Charlotte Hornets	Boston Celtics	Atlanta Hawks
Atlanta Hawks	Minnesota Timberwolves	Brooklyn Nets	Charlotte Hornets	Minnesota Timberwolves
Charlotte Hornets	Miami Heat	Cleveland Cavaliers	Los Angeles Clippers	Los Angeles Clippers
Los Angeles Clippers	Charlotte Hornets	Atlanta Hawks	San Antonio Spurs	Los Angeles Lakers
New York Knicks	San Antonio Spurs	Los Angeles Lakers	New York Knicks	Toronto Raptors
New Orleans Pelicans	Chicago Bulls	New York Knicks	Chicago Bulls	Cleveland Cavaliers

Dari hasil hasil prediksi tersebut berikut perbedaan hasil prediksi klasemen musim 2021/2022 pada Tabel 4.14.

Tabel 4.14. Perbedaan klasemen asli dan model prediksi musim lalu

Bentuk model	Metode	Perbedaan Tim yang menduduki 20 besar
Model 5	<i>Quadratic Multivariate</i>	Los Angeles Lakers, Portland Trail Blazers dan San Antonio Spurs
Model 6	<i>Quadratic Multivariate</i>	Los Angeles Lakers
Model 5	Regresi linear	Los Angeles Lakers, Portland Trail Blazers dan San Antonio Spurs
Model 6	Regresi linear	San Antonio Spurs dan Los Angeles Lakers

Setelah melakukan perbandingan antara hasil prediksi klasemen *regular season* musim 2021/2022 dengan hasil klasemen sebenarnya, keempat model di atas digunakan untuk memprediksi hasil pertandingan NBA di bulan oktober 2022 yaitu bulan pertama di musim 2022/2023. Keakuratan hasil prediksi tersebut ditampilkan pada Tabel 4.15.

Tabel 4.15. Akurasi 2 metode model

Model	Metode	Hasil akurasi model prediksi
Model 5	Metode <i>Quadratic Multivariate</i>	0.64705
Model 6	Metode <i>Quadratic Multivariate</i>	0.65686
Model 5 pembandingan	Metode Regresi Linear	0.64705
Model 6 pembandingan	Metode Regresi Linear	0.64215

Dari Tabel 4.15 dapat dilihat bentuk Model 6 metode *Quadratic Multivariate* yang paling akurat untuk memprediksi hasil pertandingan bulan Oktober di musim 2022/2023. Selanjutnya Model 6 ini digunakan untuk melakukan prediksi hingga akhir musim 2022/2023. Hasil prediksi ini ditampilkan pada gambar 4.2.

Gambar 4.2. Hasil prediksi Klasemen musim 2022/2023

team	all win	win_home	win_visitor	lose_home	all lose	lose_visitor	Semua	Menang_persen
Sacramento Kings	43	25	18	16	39	23	82	52.439024
New York Knicks	42	36	6	5	40	35	82	51.219512
Toronto Raptors	42	30	12	11	40	29	82	51.219512
Minnesota Timberwolves	42	19	23	22	40	18	82	51.219512
San Antonio Spurs	42	16	26	25	40	15	82	51.219512
Los Angeles Lakers	42	12	30	29	40	11	82	51.219512
Denver Nuggets	42	8	34	33	40	7	82	51.219512
Utah Jazz	42	4	38	37	40	3	82	51.219512
Oklahoma City Thunder	41	41	0	0	41	41	82	50.000000
Orlando Magic	41	40	1	1	41	40	82	50.000000
Detroit Pistons	41	38	3	3	41	38	82	50.000000
Houston Rockets	41	37	4	4	41	37	82	50.000000
New Orleans Pelicans	41	29	12	12	41	29	82	50.000000
Los Angeles Clippers	41	28	13	13	41	28	82	50.000000
Dallas Mavericks	41	20	21	21	41	20	82	50.000000
Miami Heat	41	20	21	21	41	20	82	50.000000
Golden State Warriors	41	13	28	28	41	13	82	50.000000
Atlanta Hawks	41	10	31	31	41	10	82	50.000000
Brooklyn Nets	41	8	33	33	41	8	82	50.000000
Memphis Grizzlies	41	6	35	35	41	6	82	50.000000
Phoenix Suns	41	1	40	40	41	1	82	50.000000
Milwaukee Bucks	41	0	41	41	41	0	82	50.000000
Washington Wizards	40	33	7	8	42	34	82	48.780488
Portland Trail Blazers	40	31	9	10	42	32	82	48.780488
Cleveland Cavaliers	40	31	9	10	42	32	82	48.780488
Chicago Bulls	40	25	15	16	42	26	82	48.780488
Indiana Pacers	40	22	18	19	42	23	82	48.780488
Philadelphia 76ers	40	13	27	28	42	14	82	48.780488
Charlotte Hornets	40	2	38	39	42	3	82	48.780488
Boston Celtics	39	17	22	24	43	19	82	47.560976



Universitas
Pertamina

Hak Cipta milik Universitas Pertamina

Dilindungi Undang-Undang

© Copyright of Universitas Pertamina



1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah kecependekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

Universitas
Pertamina



BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini mengusulkan penggunaan *Multivariate Quadratik* untuk memprediksi hasil pertandingan bola basket di liga NBA. Hasil prediksi juga dibandingkan dengan penggunaan regresi linier. Ada 6 model yang diuji untuk setiap metode. Dari semua metode tersebut, penelitian ini menunjukkan bahwa Model 6 mempunyai keakuratan terbaik diantara model lainnya. Model 6 menggunakan Dataset 22 dan hanya memprediksi *total point* tim. Tingkat keakuratan prediksi yaitu sebesar 0.84278 pada musim 2021/2022, dan 0.65686 pada bulan Oktober untuk musim 2022/2023. Selain itu, metode *Multivariate Quadratik* lebih baik dibandingkan regresi linier untuk semua model. Adapun fitur-fitur yang berpengaruh besar terhadap performa adalah sebagai berikut: total poin sebuah tim dipengaruhi oleh poin yang dicetak pemain, tembakan 3 poin yang berhasil dilakukan pemain, tembakan bebas yang berhasil dilakukan pemain, dan umpan bola terakhir dari pemain dalam satu tim sebelum mencetak poin.

5.2 Saran

Pada penelitian model prediksi ini dilakukan dengan membuat performa tiap 1 tim pada pertandingan dari data 5 pemain pada tim bola basket *NBA* dari jumlah waktu bermain paling banyak pada saat pertandingan melawan tim lawan serta diurutkan sesuai posisi dalam satu baris dan fitur fitur pada tiap pemain yang mempengaruhi hasil pertandingan tersebut. Saran untuk penelitian kedepan dapat melakukan prediksi pada pembuatan model data pada satu baris performa 5 pemain tim dan tim lawan dengan pada saat pertandingan mengambil waktu bermain paling banyak serta mengurutkan dari posisi di jadikan 1 baris antara tim dan tim lawan dan dengan *features selection*, dan membuat model sama atau menggunakan metode lainnya .



Universitas
Pertamina

Hak Cipta milik Universitas Pertamina

Dilindungi Undang-Undang

© Copyright of Universitas Pertamina



1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah kecependekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

Universitas
Pertamina

DAFTAR PUSTAKA

- 2playbook (2021). Acb: renovación de endesa hasta 2024 y nuevos espónsors para sostener un negocio de 30 millones), Fact Sheet N°282. https://www-2playbook-com.translate.google.com/competiciones/acb-cierra-ano-pandemia-con-ingresos-record-302-millones-euros_3657_102.html?_x_tr_sl=auto&_x_tr_tl=en&_x_tr_hl=id&_x_tr_pto=wapp, Last accessed on 2022-12-01.
- Azubuike, I. M. (2019). Second order regression with two predictor variables centered on mean in an conditioned model. *International Journal of Statistics and Applications*, 9(4):101–110.
- Battista and Alojsio, S. (2019). *The NBA's economic environmet: how a professional basketball league managed to bounce back from a lockout situation*. Luiss Guido Carli.
- Chibani, S. and Coudert, F.-X. (2020). Machine learning approaches for the prediction of materials properties. *Apl Materials*, 8(8):080701.
- Cholissodin and Imam (2019). *Buku Ajar AI, Machine Learning & Deep Learning*. Malang: Fakultas Ilmu Komputer (FILKOM) Universitas Brawijaya (UB).
- Corella Parra, B. A. (2021). Using big data techniques to measure the performance of professional basketball teams. B.S. thesis, Universidad de Investigación de Tecnología Experimental Yachay.
- Dimitrije Curcic (2021). Nba revenue statistics (2001-2022), Fact Sheet N°282. <https://runrepeat.com/nba-revenue-statistics>, Last accessed on 2012-12-01.
- Hsu, P.-H., Galsanbadam, S., Yang, J.-S., and Yang, C.-Y. (2018). Evaluating machine learning varieties for nba players' winning contribution. In *2018 International Conference on System Science and Engineering (ICSSE)*, pages 1–6. IEEE.
- Id, I. D. (2021). *Machine Learning: Teori, Studi Kasus dan Implementasi Menggunakan Python*, volume 1. Pekanbaru : Universitas Riau.
- Imran, H., Al-Abdaly, N. M., Shamsa, M. H., Shatnawi, A., Ibrahim, M., and Ostrowski, K. A. (2022). Development of prediction model to predict the compressive strength of eco-friendly concrete using multivariate polynomial regression combined with stepwise method. *Materials*, 15(1):317.
- Jebli, I., Belouadha, F.-Z., Kabbaj, M. I., and Tilioua, A. (2021). Prediction of solar energy guided by pearson correlation using machine learning. *Energy*, 224:120109.
- Nguyen, N. H., Nguyen, D. T. A., Ma, B., and Hu, J. (2022). The application of machine learning and deep learning in sport: predicting nba players' performance and popularity. *Journal of Information and Telecommunication*, 6(2):217–235.
- Oytun, M., Tinazci, C., Sekeroglu, B., Acikada, C., and Yavuz, H. U. (2020). Performance prediction and evaluation in female handball players using machine learning models. *IEEE Access*, 8:116321–116335.
- Severini, T. A. (2020). *Analytic methods in sports: Using mathematics and statistics to understand data from baseball, football, basketball, and other sports*. Chapman and Hall/CRC.
- Sinha, P. (2013). Multivariate polynomial regression in data mining: methodology, problems and solutions. *Int. J. Sci. Eng. Res*, 4(12):962–965.

Stathopoulos, A. and Karlaftis, M. G. (2003). A multivariate state space approach for urban traffic flow modeling and prediction. *Transportation Research Part C: Emerging Technologies*, 11(2):121–135.

Uğur Yılmaz (2022). European basketball revenue), Fact Sheet N°282. <https://chroniclingbasketball.net/tag/european-basketball-revenue/>, Last accessed on 2022-12-01.

Young, B. W. and Starkes, J. L. (2005). Career-span analyses of track performance: longitudinal data present a more optimistic view of age-related performance decline. *Experimental Aging Research*, 31(1):69–90.

1. Dilarang mengutip karya tulis ini, kecuali:

- a. menyebutkan sumber sesuai kaidah kecependekiaan;
- b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
- c. pengutipan tidak merugikan Universitas Pertamina.

2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.



Universitas
Pertamina

Hak Cipta milik Universitas Pertamina

Dilindungi Undang-Undang

© Copyright of Universitas Pertamina



1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah kecependekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

Universitas
Pertamina



Form TA-2 Bimbingan Tugas Akhir
FAKULTAS SAINS DAN ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER

Nama Mahasiswa : Boy Cakraningrat

NIM : 105217028

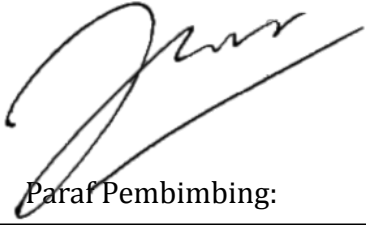
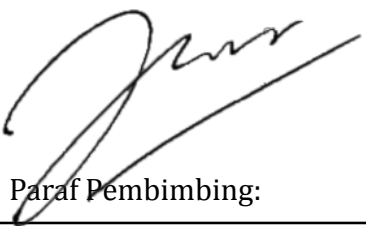
Nama Pembimbing :Rangga Ganzar Noegraha, Ph.D

NIP :119005

No.	Hari/Tanggal: 30-09-2022
<p>Hal yang menjadi perhatian:</p> <p>Penjelasan Topic penelitian tentang nba dan metode pengumpulan data</p> <p>Paraf Pembimbing:</p>	
No.	Hari/Tanggal: 19-10-2022
<p>Hal yang menjadi perhatian:</p> <p>Hasil dari pengumpulan data dan melakukan pre processing</p> <p>Paraf Pembimbing:</p>	

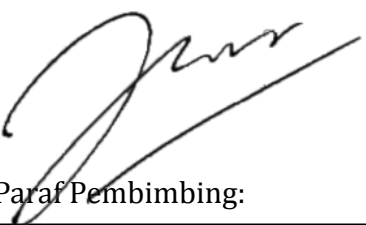
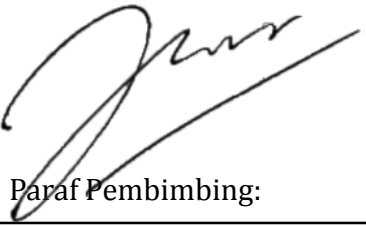


Form TA-2 Bimbingan Tugas Akhir
FAKULTAS SAINS DAN ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER

No.	Hari/Tanggal: 26-10-2022
<p>Hal yang menjadi perhatian:</p> <p>melakukan pre processing penelitian tentang nba dan metode fitur seleksi</p> <p style="text-align: right;"> Paraf Pembimbing:</p>	
No.	Hari/Tanggal: 2-11-2022
<p>Hal yang menjadi perhatian:</p> <p>Hasil dari pengumpulan data dan melakukan pre processing</p> <p style="text-align: right;"> Paraf Pembimbing:</p>	



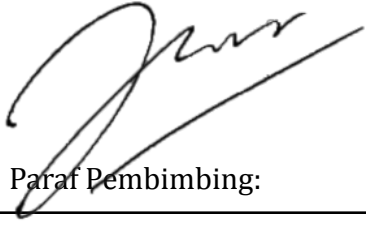
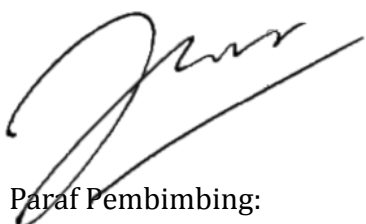
Form TA-2 Bimbingan Tugas Akhir
FAKULTAS SAINS DAN ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER

No.	Hari/Tanggal: 10-11-2022
<p>Hal yang menjadi perhatian:</p> <p>Penyusunan bab 1 dan 2</p> <p style="text-align: right;">  Paraf Pembimbing: </p>	
No.	Hari/Tanggal: 24-12-2022
<p>Hal yang menjadi perhatian:</p> <p>penyusunan bab 3</p> <p style="text-align: right;">  Paraf Pembimbing: </p>	

2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

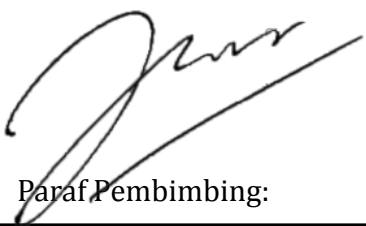
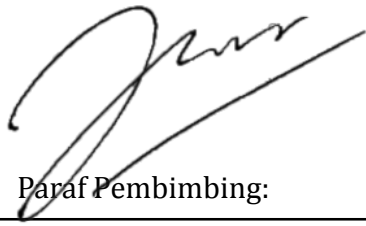


Form TA-2 Bimbingan Tugas Akhir
FAKULTAS SAINS DAN ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER

No.	Hari/Tanggal: 2-12-2022
Hal yang menjadi perhatian: penyusunan laporan dan seminar kemajuan	
 Paraf Pembimbing:	
No.	Hari/Tanggal: 31-1-2023
Hal yang menjadi perhatian: pembuatan model prediksi	
 Paraf Pembimbing:	



Form TA-2 Bimbingan Tugas Akhir
FAKULTAS SAINS DAN ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER

No.	Hari/Tanggal: 1-02-2023
<p>Hal yang menjadi perhatian:</p> <p>Pembuatan model prediksi multivariate</p> <p></p> <p>Paraf Pembimbing:</p>	
No.	Hari/Tanggal: 07-02-2023
<p>Hal yang menjadi perhatian:</p> <p>pembuatan model prediksi musim ini</p> <p></p> <p>Paraf Pembimbing:</p>	



Form TA-2 Bimbingan Tugas Akhir
**FAKULTAS SAINS DAN ILMU
KOMPUTER PROGRAM STUDI ILMU
KOMPUTER**

Nama Mahasiswa :Boy Cakraningrat

NIM :105217028

Nama Pembimbing :Ade Irawan,Ph.D

NIP :116130

No. Hari/Tanggal: 01- 12-2022

Hal yang menjadi perhatian:

Memberikan draf awal bab 1 dan bab 2

Paraf Pembimbing:

No. Hari/Tanggal: 12/12/2022

Hal yang menjadi perhatian:

Revisi draft seminar

Paraf Pembimbing:



Form TA-2 Bimbingan Tugas Akhir
**FAKULTAS SAINS DAN ILMU
KOMPUTER PROGRAM STUDI ILMU
KOMPUTER**

No.

Hari/Tanggal: 04- 01-2023

Hal yang menjadi perhatian:

membicarakan outlier
bentuk model prediksi dan feature selection

Paraf Pembimbing:

No.

Hari/Tanggal: 31/01/2023

Hal yang menjadi perhatian:

membuat model prediksi multivariate

Paraf Pembimbing:



Form TA-2 Bimbingan Tugas Akhir
**FAKULTAS SAINS DAN ILMU
KOMPUTER PROGRAM STUDI ILMU
KOMPUTER**

No.

Hari/Tanggal: 01- 02-2023

Hal yang menjadi perhatian:

bentuk model prediksi dan feature
selection dengan model asli

Paraf Pembimbing:

No.

Hari/Tanggal: 02/02/2023

Hal yang menjadi perhatian:

latex membenarkan bentuk tabel
hasil dari prediksi musim ini

Paraf Pembimbing:



Form TA-2 Bimbingan Tugas Akhir
**FAKULTAS SAINS DAN ILMU
KOMPUTER PROGRAM STUDI ILMU
KOMPUTER**

No.

Hari/Tanggal: 03- 02-2023

Hal yang menjadi perhatian:

membicarakan outlier
membuat model pembandingan

Paraf Pembimbing:

No.

Hari/Tanggal: 06/02/2023

Hal yang menjadi perhatian:

membandingkan hasil model pembandingan
dengan model multivariate

Paraf Pembimbing:



Universitas
Pertamina

Hak Cipta milik Universitas Pertamina

Dilindungi Undang-Undang

© Copyright of Universitas Pertamina

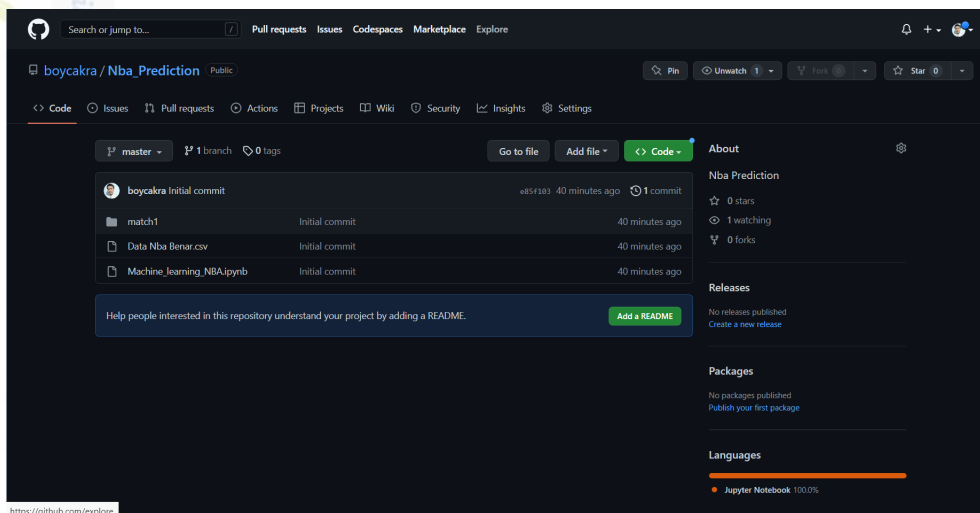


1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah kecependekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

Universitas
Pertamina

LAMPIRAN A

Kumpulan data mentah nba season 2021-2022 dan code



Gambar A.1. Data kumpulan data

2.
 - a. menyebutkan sumber sesuai kaidah kecendekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.

<https://github.com/explore>