

Research Article

Analisis Ramalan Siri Masa Hujan menggunakan Aplikasi ramalHujan

Azlarul Ikhmal Hakim Mohd Puad¹, Zuraini Zainol^{2,*}, dan Ummul Fahri Abdul Rauf³

¹ Jabatan Sains Komputer, Universiti Pertahanan Nasional Malaysia; azlarulhakim@gmail.com;

² Jabatan Sains Komputer, Universiti Pertahanan Nasional Malaysia; zuraini@upnm.edu.my;

 ORCID ID 0000-0002-6881-7039

³ Pusat Asasi Pertahanan, Universiti Pertahanan Nasional Malaysia; ummul@upnm.edu.my;

 ORCID ID 0000-0001-6080-528X

* Correspondence: zuraini@upnm.edu.my

Abstrak: Ramalan hujan memainkan peranan penting dalam perancangan strategik dan pengurusan risiko, terutamanya dalam menghadapi cabaran perubahan iklim dan cuaca ekstrem. Kajian ini menggunakan pendekatan Seasonal Autoregressive Integrated Moving Average (SARIMA) bagi menganalisis ramalan data hujan bulanan di stesen pengumpulan data hujan di Sungai Manggis, Selangor. Data hujan bulanan dalam tempoh Jan 1970 hingga Disember 2014 diperolehi daripada Jabatan Pengairan dan Saliran digunakan sebagai asas untuk membangunkan model ramalan hujan. Proses pembangunan model melibatkan pembersihan data, penggantian nilai hilang, dan analisis pola data bermusim. Ketepatan model dinilai menggunakan Akaike Information Criterion (AIC) dan unjuran ramalan sehingga lima tahun ke hadapan digunakan untuk mengenal pasti pola hujan yang konsisten. Model ramalan yang dibangunkan ini diintegrasikan bersama aplikasi ramalHujan, yang membolehkan pengguna membuat ramalan jumlah hujan berdasarkan lokasi dan bilangan bulan yang ditentukan. Keputusan ramalan hujan divisualisasikan dalam bentuk carta interaktif untuk memudahkan analisis lebih lanjut. Hasil kajian menunjukkan bahawa model SARIMA memberikan prestasi yang sangat baik dalam mengenal pasti trend bermusim, dan aplikasi ramalan yang dihasilkan membolehkan ramalan jangka masa pendek yang lebih tepat. Kajian ini memberikan sumbangan penting kepada pengurusan sumber air dan mitigasi bencana, dengan potensi aplikasi yang luas untuk perancangan dan pengurusan risiko yang lebih berkesan.

Kata kunci: ramalan hujan;; analisis siri masa; perubahan iklim; SARIMA.

DOI: 10.5281/zenodo.1482225



Copyright: © 2025 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. PENGENALAN

Hujan merupakan salah satu komponen penting dalam kitaran air (Wang & Xie, 2018) yang mempengaruhi pelbagai bidang seperti pertanian, hidrologi, pengurusan sumber air, mitigasi bencana (Loh et al., 2024), dsb. Walau bagaimanapun, kesan perubahan iklim semakin dirasai di Malaysia seperti hujan yang tidak menentu, telah memberi kesan buruk kepada sektor-sektor ini. Oleh itu, ketepatan ramalan hujan adalah kritikal dalam memastikan keberkesanan pengurusan sektor-sektor ini. Namun, faktor perubahan iklim yang tidak menentu serta faktor-faktor hujan seperti angin, kelembapan dan suhu telah menjadikan ramalan hujan satu cabaran yang besar (Rahman et al., 2022; Zainudin et al., 2016). Bagi meningkatkan ketepatan ramalan hujan, pelbagai teknik telah digunakan termasuklah kaedah statistik, perlombongan data, pembelajaran mesin, pengkomputeran awan dan data raya.

Melalui teknik-teknik ini, penyelidik dapat membina model siri masa yang mampu mengenal pasti corak hujan berdasarkan data sejarah (*historical data*) curahan hujan, dan seterusnya memberi ramalan yang lebih tepat bagi kejadian hujan masa hadapan.

Menurut Han et al. (2011), perlombongan data (PD) ialah proses pengekstrakan data dan corak tersembunyi di dalam data berskala besar menggunakan algoritma tertentu. Secara umumnya, PD merangkumi pelbagai algoritma untuk menyelesaikan tugas tertentu seperti pengelasan, regresi, analisis siri masa, ramalan, pengelompokan, dsb. (Dunham, 2006). Ringkasan berkenaan model, tugas dan algoritma dalam PD boleh dirujuk di dalam kajian Mutalib et al. (2023) dan Nohuddin et al. (2018b). Kombinasi teknologi, peranti dan pelbagai teknik seperti PD, pembelajaran mesin dan *Internet of Things* (IOT) telah digunakan secara meluas untuk mengkaji pelbagai aspek perubahan iklim. Antaranya ialah ramalan dan analisis corak hujan (Lim et al., 2024; Rashid et al., 2017, 2018; Tee & Mansor, 2024), ramalan bencana alam seperti banjir (Marzukhi et al., 2018; Oddo et al., 2024; Vun et al., 2024), pengoptimuman tenaga boleh diperbaharui (Bassey, 2023; Nohuddin et al., 2018a; Yao et al., 2023), serta kesan perubahan iklim terhadap sektor pertanian (Aworka et al., 2022; Kalimuthu et al., 2020), dsb. Salah satu tugas yang sering digunakan untuk tujuan ramalan hujan ialah analisis siri masa. Teknik ini membantu dalam mengenal pasti *trend* dan pola berdasarkan data sejarah untuk memberikan hasil ramalan yang lebih tepat.

Menurut Dunham (2006), siri masa merujuk kepada satu set urutan tempoh masa yang memegang atribut dengan nilai tertentu. Secara ringkasnya, siri masa merupakan kumpulan data yang direkodkan mengikut susunan masa semasa data tersebut dikumpulkan (Amin et al., 2014; Nohuddin et al., 2018c). Antara aplikasi PD siri masa yang sering digunakan adalah mengenal pasti persamaan antara dua siri masa yang berbeza serta meramalkan nilai masa depan bagi suatu atribut tertentu. Dalam konteks ramalan atribut yang bergantung pada siri masa, analisis awal diperlukan untuk memahami corak data. Analisis siri masa melibatkan pencarian corak data dan peramalan nilai masa depan dengan memberi tumpuan kepada empat aspek utama iaitu *trend*, kitaran (*cycles*), bermusim (*seasonal*) dan *outliers*. Untuk menganalisis *trend*, kaedah pelicinan (*smoothing*) merupakan pendekatan yang efektif bagi mengurangkan *nonsystematic behavior* yang terdapat dalam siri masa. Kebiasaannya, kaedah pelicinan menggunakan pendekatan seperti purata pergerakan (*moving average*) untuk atribut yang dianalisis. Tujuan utama teknik ini ialah untuk untuk menapis hingar (*noise*) dan outliers.

Ramalan data siri masa merupakan pendekatan yang sering terdedah kepada ralat dan kehadiran hingar. Salah satu kaedah untuk meramalkan siri masa masa hadapan adalah melalui teknik *autoregression*, yang menggunakan nilai data daripada siri masa sebelumnya sebagai asas ramalan. Teknik ini dapat diterapkan pada siri masa pegun (*stationary*) atau bukan pegun (*nonstationary*), bergantung pada sifat data yang dianalisis. Apabila kaedah *autoregression* digabungkan dengan *moving average*, ia membentuk model ramalan siri masa yang dikenali sebagai *Autoregressive Moving Average* (ARMA). Model ARIMA terhasil daripada penambahbaikan model ARMA untuk menangani masalah data siri masa yang tidak pegun (Majka, 2024). ARIMA adalah model statistik linear yang digunakan untuk meramalkan nilai sesuatu peristiwa berdasarkan data dalam satu tempoh masa (Lai & Dzombak, 2020; Mishra & Desai, 2005). Model ini dibangunkan untuk siri masa pegun dan melibatkan proses *differencing* (mengira perbezaan antara pemerhatian berturutan) untuk mengubah siri bukan pegun menjadi pegun. Algoritma ini dilatih dengan menganalisis pola data masa lampau dan digunakan untuk meramalkan nilai masa ke hadapan dalam satu siri. Bagi siri masa pegun yang mempunyai pola bermusim, model *Seasonal Autoregressive Integrated Moving Average* (SARIMA) lebih sesuai digunakan berbanding ARIMA kerana kemampuannya untuk menunjukkan fluktuasi (*fluctuation*) berkala dengan lebih teratur (Majka, 2024).

Objektif utama kajian ini adalah (i) Mengumpul set data hujan dengan atribut yang diperlukan dalam kajian penyelidikan ini; (ii) Membina model ramalan data hujan bulanan menggunakan model

SARIMA, dan (iii) Membangunkan satu aplikasi yang dapat memvisualisasikan keputusan ramalan hujan.

2. KAJIAN LITERATUR

Beberapa kajian telah dijalankan untuk meneliti perubahan iklim dan meramal data hujan di Malaysia. Tee and Mansor (2024) dalam kajian mereka menggunakan model siri masa univariat untuk meramalkan hujan di Selangor berdasarkan data dari stesen Petaling Jaya dan Subang. Model ARIMA (2,0,3) dan ARIMA (4,0,4) dikenal pasti sebagai model terbaik bagi kedua-dua stesen tersebut. Hasil kajian mereka meramalkan puncak hujan pada bulan Mei dan Jun 2023 dengan curahan hujan yang stabil sehingga akhir tahun.

Poh and Norrulashikin (2022) membangunkan model ramalan hujan bulanan di Semenanjung Malaysia menggunakan model ARIMA dan *Holt-Winter*. Kajian ini menggunakan data hujan dari Januari 2010 sehingga Disember 2018 di enam stesen utama, iaitu Senai, Pulau Langkawi, Kula Terengganu, Gong Kedak, Cameron highlands dan Alor Star sebagai set data latihan. Manakala data ujian bermula dari Januari 2019 sehingga September 2019 digunakan untuk menguji ketepatan ramalan model yang telah dibangunkan. Hasil kajian mereka mendapati model ARIMA memberikan prestasi yang lebih baik berbanding kaedah Holt-Winter.

Dalam kajian lain, Samsulaiman and Ismail (2023) menggunakan teknik SARIMA dan kaedah *Holt-Winters's Exponential Smoothing* untuk membangunkan model ramalan hujan. Dua stesen di Terengganu dan Pahang telah dipilih sebagai kawasan kajian, dan data hujan bulanan dari Januari 1980 hingga Disember 2021 diperolehi daripada Jabatan Pengairan dan Saliran Malaysia dan Jabatan Meteorologi Malaysia. Dapatan kajian menunjukkan bahawa model SARIMA (0,1,1)(0,1,1)₁₂ (stesen Kg. Dura), SARIMA (0,1,1)(0,1,1)₁₂ (stesen Al-Muktafi Billah Shah), SARIMA (0,1,1)(1,0,1)₁₂ (stesen Sg. Lembing PCCL Mill) dan SARIMA (0,1,1)(0,0,1)₁₂ (Kg. Sg. Yap). Hasil kajian juga mendapati kaedah *Holt-Winters's Exponential Smoothing* memberikan keputusan yang lebih tepat dan sesuai digunakan untuk meramal hujan pada masa hadapan.

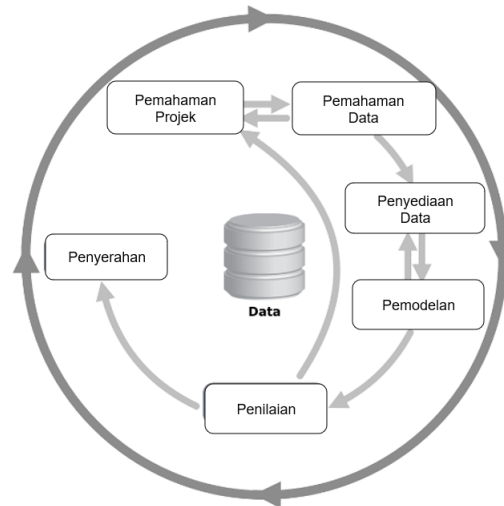
Mustapha and Ismail (2021) telah membandingkan keberkesanan dua model analisis siri masa univariat, iaitu *Autoregressive Integrated Moving Average* (ARIMA) dan SARIMA, dalam meramalkan siri masa hujan bulanan di Kelantan, Malaysia. Hasil kajian menunjukkan bahawa model SARIMA merupakan model yang lebih sesuai untuk meramalkan siri masa hujan bulanan, kerana ia menghasilkan ralat pengukuran yang lebih rendah.

Berdasarkan pemerhatian di atas, didapati semua kajian membangunkan model ramalan hujan menggunakan data daripada beberapa stesen pilihan bagi menganalisis perubahan dalam iklim. Selain itu, model ARIMA dan SARIMA dikenal pasti sebagai kaedah yang paling kerap digunakan untuk meramal hujan di Malaysia.

3. METODOLOGI KAJIAN

Kajian ini menggunakan metodologi *Cross Industry Standard Process for Data Mining* (CRISP-DM) yang mengandungi 6 fasa utama iaitu pemahaman projek, pengumpulan data, penyediaan data, pemodelan analisis, penilaian dan penyerahan. Fasa 1 melibatkan pemahaman kajian dan merancang keperluan penyelidikan. Seperti kebanyakan sistem analisis, langkah pertama dalam proses analisis memerlukan penulis memahami objektif dan keperluan yang hendak dicapai daripada kajian ini. Kajian ini memfokuskan kepada pembangunan model ramalan menggunakan kaedah pembelajaran

mesin berdasarkan siri masa. Pemahaman yang berkaitan dengan analisis siri masa diterapkan untuk membangunkan model SARIMA. Model ini dibangunkan berdasarkan data taburan hujan bulanan yang diperolehi daripada Jabatan Pengairan dan Saliran (JPS) di stesen pengumpulan data di negeri Selangor iaitu Sungai Manggis dalam tempoh 44 tahun yang bermula daripada Januari 1970 sehingga Disember 2014.



Rajah 1. CRISP-DM diadaptasi daripada (Wirth & Hipp, 2000).

Fasa 2 memberi tumpuan kepada pengumpulan set data. Kajian ini menggunakan set data yang diperolehi daripada JPS yang mengandungi keseluruhan negeri di Malaysia. Set data disimpan dalam format teks (*.txt) dan diasingkan mengikut set data harian dan bulanan. Dalam kajian ini, data hujan yang digunakan terdiri daripada jumlah hujan harian yang kemudiannya dikategorikan sebagai data hujan bulanan bermula dari tahun 1970 sehingga 2014 di semua stesen di Selangor. Untuk kajian ini, stesen Sungai Manggis telah dipilih sebagai lokasi kajian. Data mentah ini akan melalui proses transformasi menggunakan kaedah pencairan. Penyediaan data merupakan langkah penting dalam memastikan set data yang bersih tersedia untuk digunakan pada fasa pemodelan berikutnya. Dalam fasa ini, proses pemilihan atribut, pembersihan data, dan transformasi data akan dilakukan secara menyeluruh. Set data akan dianalisis untuk mengenal pasti masalah seperti data yang tidak bersih, data hilang, dan data tidak konsisten. Tujuan utama pra-pemprosesan adalah untuk membersihkan set data bagi memastikan hasil ramalan akhir adalah tepat dan berkualiti (Noormanshah et al., 2020; Zainol et al., 2018). Sekiranya data tidak diproses dengan baik, ia boleh menyebabkan ketidaktepatan dalam ramalan (Zainol et al., 2019). Rajah 2 menunjukkan contoh 21 baris data mentah yang diekstrak daripada tolok hujan di stesen Pejabat JPS Sungai Manggis, Selangor.

1	~~~ NIWA Tideda ~~~ JPS Ampang													19-JUL-2016 11:04
2	~~~ PCAL ~~~													
3														
4	Source is Z:\MIX\hujan selangor.mtd													
5	Monthly totals 1970 to 2014													site 2815001 PEJABAT JPS. SG. MANGG at SELANGOR
6														
7	Rain mm													
8														
9	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
10														
11	1970	?	32.7	170.5	209.3	151.7	134.2	147.8	114.3	379.1	345.2	226.8	242.8	2154.3?
12	1971	280.5	84.9	143.6	87.1	50.6	44.9	119.4	349.7	202.0	156.5	109.3	181.4	1809.9
13	1972	96.2	164.4	61.5	178.3	63.0	99.9	127.7	120.7	206.9	157.0	177.7	353.7	1807.1
14	1973	189.5	268.3	160.0	188.4	246.3	109.2	142.1	164.5	182.3	216.7	376.5	209.5	2453.2
15	1974	51.0	80.5	45.0	170.0	128.0	12.5	137.0	67.0	365.0	109.5	137.0	151.5	1454.0
16														
17	1975	163.0	104.0	136.0	189.5	96.5	78.0	68.5	121.5	155.0	159.0	234.5	38.0	1543.5
18	1976	133.0	27.5	182.5	198.0	66.0	161.5	58.0	180.0	136.0	265.0	243.5	185.0	1836.0
19	1977	72.5	72.5	22.5	73.5	123.0	133.0	74.5	141.5	175.0	145.5	361.0	210.0	1604.5
20	1978	195.5	142.0	167.0	76.0	77.0	51.0	97.0	253.0	172.5	208.5	288.5	151.0	1879.0
21	1979	176.0	107.0	40.0	139.0	111.0	224.5	238.0	186.0	101.5	251.0	223.0	137.5	1934.5
22														

Rajah 2. Contoh set data mentah untuk stesen Pejabat JPS Sungai Manggis, Selangor.

Set data (rujuk Rajah 2) mengandungi kehilangan data pada baris 11 di lajur bulan Jan, yang ditandakan dengan simbol tanda soal '?'. Untuk mengendalikan data hilang, penulis menggunakan kaedah penggantian nilai dengan purata rekod data bagi tiga bulan sebelumnya. Pendekatan ini dilakukan dengan memanfaatkan sintaks Python, iaitu *Pandas fillna()*, sebagai arahan untuk mengisi bahagian data yang hilang. Proses ini digabungkan dengan penggunaan sintaks *rolling()* dan *mean()*. Fungsi *rolling()* ditetapkan kepada tettingkap berukuran 3, yang merujuk kepada tiga sel data sebelumnya, sementara parameter *min_periods* digunakan untuk memastikan nilai purata dikira hanya jika data yang mencukupi tersedia. Dengan cara ini, nilai purata tiga bulan sebelumnya bagi set data yang hilang dapat diperoleh dan dimasukkan secara automatik. Rajah 3 menunjukkan paparan kod Python yang lebih terperinci mengenai proses ini.

```
#Fillna Dataset 1

df1 = df3.fillna(df3.rolling(3,min_periods=1).mean())
df1.fillna(0, inplace=True)

#Fillna Dataset 2

df2 = df2.fillna(df2.rolling(3,min_periods=1).mean())
df2.fillna(0, inplace=True)

#Fillna Dataset 3

df3 = df3.fillna(df3.rolling(3,min_periods=1).mean())
df3.fillna(0, inplace=True)
```

Rajah 3. Paparan kod Python untuk baris yang kehilangan data menggunakan teknik *Moving Average*

Dalam fasa pemodelan, proses pembangunan model SARIMA yang sesuai dengan set data hujan akan diterangkan sebagai langkah utama dalam membangunkan aplikasi ramalan hujan menggunakan model SARIMA. Bagi tujuan ini, penulis menggunakan pakej *statsmodels* yang tersedia dalam Python untuk menilai dan memilih pemodelan yang paling sesuai bagi tujuan ramalan hujan. Kajian ini menggunakan data hujan dari Januari 1970 sehingga Disember 2013 sebagai set data latihan. Manakala data ujian bermula dari Januari 2014 sehingga Disember 2014 digunakan untuk menguji ketepatan ramalan. Pelbagai algoritma model telah dibandingkan untuk mengenal pasti model yang paling sesuai berdasarkan sifat data. Hasil kajian ini menunjukkan model SARIMA(0, 0, 1)(0, 1, 0)₁₂ menghasilkan unjuran nilai yang terbaik dan paling tepat untuk set data hujan ini. Model ini memberikan nilai *Akaike Information Criterion* (AIC) yang berada dalam julat antara 1400 hingga 1600. Nilai ini menonjolkan kualiti relatif model SARIMA bagi menganalisis dan meramalkan hujan berdasarkan set data yang digunakan.

Fasa penilaian merupakan aspek yang penting dalam memastikan keberkesanan model yang telah dibangunkan serta ketetapan data yang digunakan. Dalam fasa ini, penulis akan merujuk kembali kepada proses dalam fasa-fasanya untuk mengenal pasti sebarang penambahbaikan atau elemen tambahan yang perlu dilakukan pada kedua-dua fasa. Tujuan utama fasa ini adalah untuk memastikan bahawa data dan model yang dibangunkan berada dalam keadaan yang optimum, bagi menghasilkan keputusan ramalan yang tepat dan boleh dipercayai. Set data hujan digunakan untuk menguji ketepatan model ramalan yang dibangunkan. Proses pengujian ini bertujuan memastikan bahawa ketepatan model memenuhi keperluan analisis yang telah ditetapkan.

Hasil daripada fasa penilaian ini adalah laporan yang perlu disediakan dan dipersembahkan kepada pengguna akhir. Dalam peringkat akhir ini, model yang dibangunkan berdasarkan siri masa ini akan dimuatkan ke dalam aplikasi ramalan. Aplikasi ini direka untuk meramal hujan berdasarkan lokasi (set data hujan berdasarkan lokasi yang dinyatakan dalam fasa 2), bilangan bulan (bulan yang ingin diramalkan). Setiap ramalan yang dibuat akan dipaparkan dengan nilai purata hujan untuk

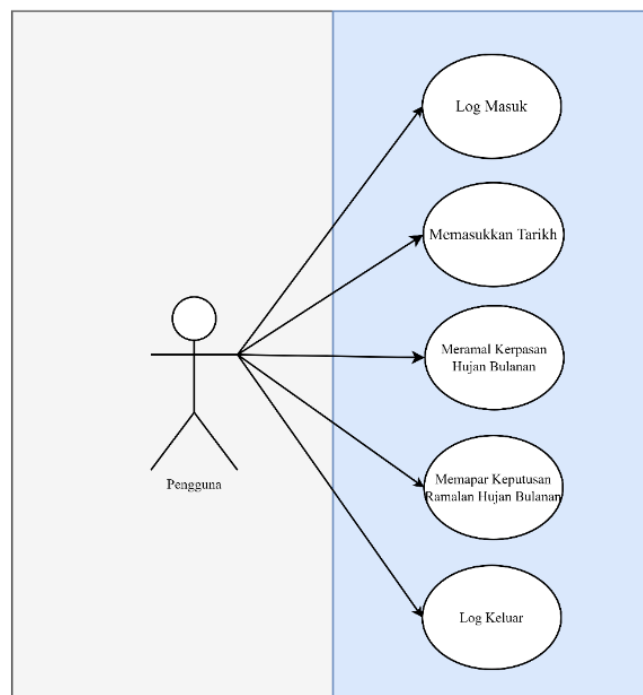
setiap bulan yang diramalkan. Akhir sekali, keputusan ramalan hujan akan divisualisasikan dalam bentuk carta yang disediakan dalam aplikasi *ramalHujan*, yang bertujuan untuk memudahkan pemahaman pengguna.

4. ANALISIS APLIKASI

Analisis aplikasi berperanan penting dalam membantu penyelidik mengenal pasti objektif, merumuskan masalah, dan mengenal pasti kekurangan pada peringkat awal pembangunan aplikasi. Penggunaan rajah seperti *use case* dan carta alir digunakan sebagai alat visual untuk menerangkan konsep dan fungsi aplikasi kepada pengguna dengan lebih jelas sepanjang proses analisis.

4.1 Gambarajah Use Case

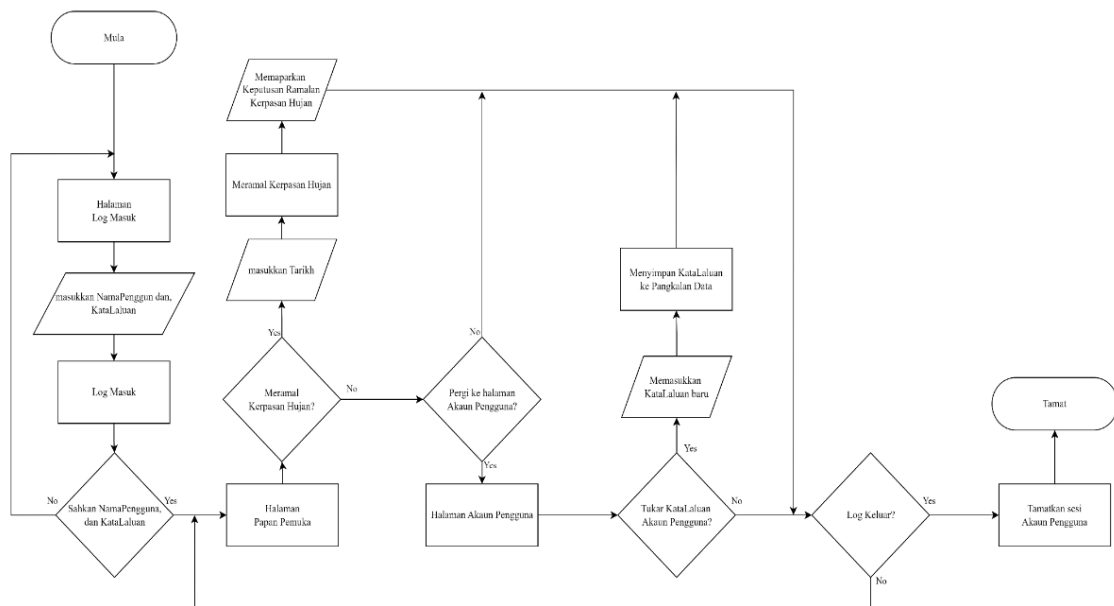
Rajah 4 menggambarkan *use case* bagi aplikasi *ramalHujan*, yang menerangkan fungsi-fungsi asas yang disediakan kepada pengguna. Antara fungsi utama aplikasi ini ialah: (i) Log Masuk: Pengguna boleh log masuk ke aplikasi untuk mengakses fungsi-fungsi lain yang tersedia. (ii) Log Keluar: Fungsi ini membolehkan pengguna menamatkan sesi mereka dengan selamat. (iii) Memasukkan Tarikh: Pengguna boleh menetapkan tarikh tertentu untuk digunakan dalam proses ramalan. (iv) Meramal Kerpasan Hujan Bulanan: Berdasarkan tarikh akhir data yang tersedia dan tarikh yang dimasukkan oleh pengguna, aplikasi ini akan membuat ramalan hujan bulanan, dan (v) Memaparkan Keputusan Ramalan: Keputusan ramalan akan dipaparkan dalam bentuk rajah visual interaktif yang memudahkan pengguna memahami pola ramalan hujan yang dihasilkan. Pendekatan *use case* ini memastikan aplikasi memberikan pengalaman pengguna yang mesra dan memenuhi keperluan mereka dalam menganalisis dan meramal data hujan dengan berkesan.



Rajah 4. Gambarajah *use case* bagi aplikasi *ramalHujan*.

4.2 Carta Alir

Rajah 5 menggambarkan carta alir bagi aplikasi *ramalHujan*, yang merincikan proses interaksi pengguna dengan aplikasi. Langkah pertama melibatkan pengguna mengakses halaman log masuk, di mana mereka perlu memasukkan *NamaPengguna* dan *KataLaluan*. Sekiranya pengguna belum memiliki akaun, mereka diberi pilihan untuk mendaftarkan akaun baharu. Selepas pendaftaran atau log masuk, pengesahan akaun pengguna dilakukan melalui pangkalan data. Sekiranya berjaya, pengguna akan dibawa ke paparan papan pemuka utama. Selain itu, aplikasi ini membolehkan pengguna menavigasi antara pelbagai halaman, termasuk halaman akaun pengguna. Setiap navigasi melibatkan syarat tertentu (*condition*) yang perlu dipenuhi untuk memastikan capaian pengguna adalah sah dan paparan halaman yang sesuai kepada pengguna. Pendekatan ini memastikan proses navigasi yang lancar dan meningkatkan pengalaman pengguna ketika berinteraksi dengan aplikasi.



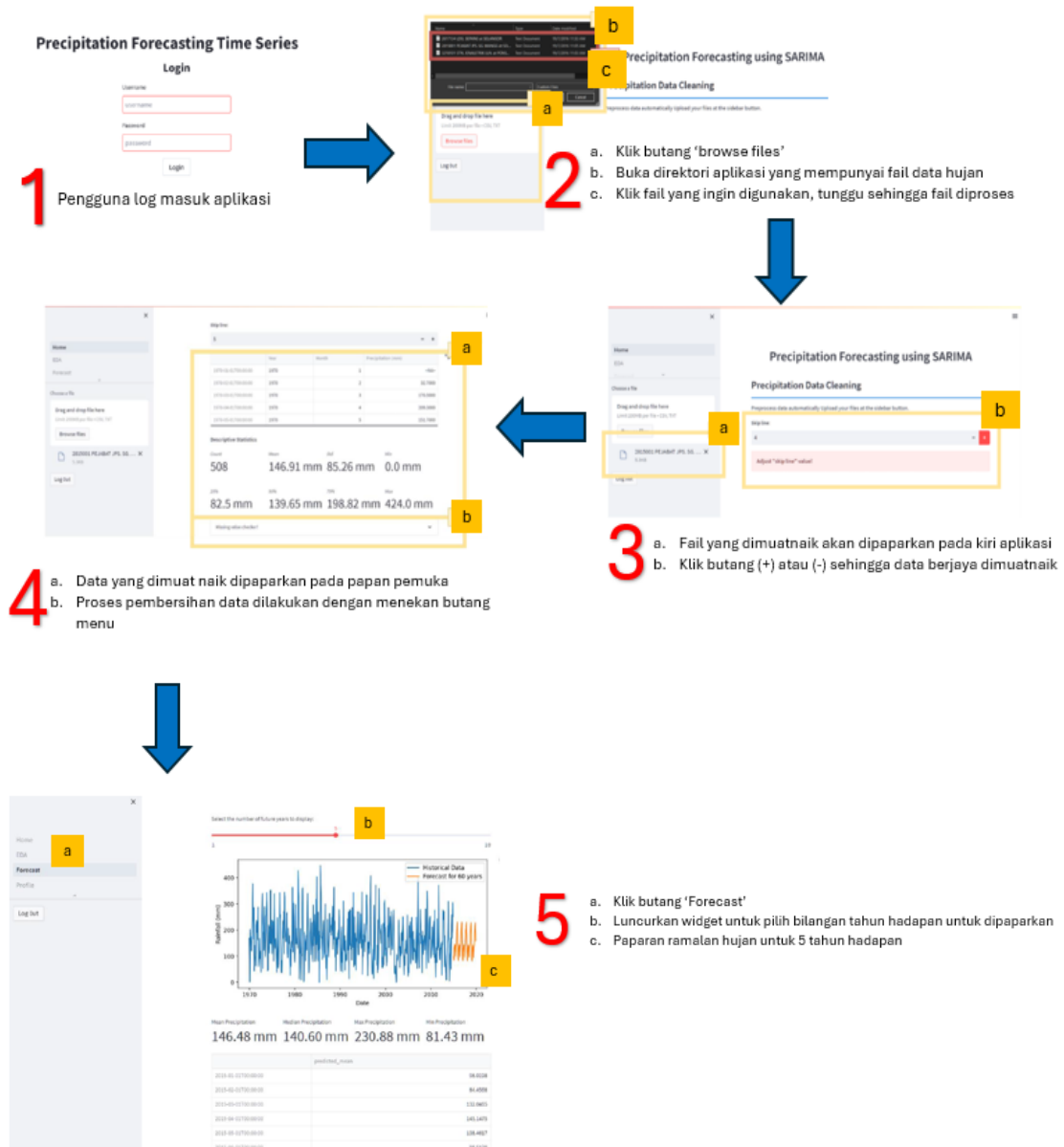
Rajah 5. Carta Alir bagi aplikasi *ramalHujan*.

5. MODEL RAMALAN DAN APLIKASI RAMALHUJAN

Model ramalan yang diperolehi di fasa pemodelan akan integrasikan dengan aplikasi *ramalHujan*. Aplikasi ini dibangunkan dengan menggunakan bahasa pengaturcaraan Python dan modul Streamlit. Streamlit merupakan rangka kerja sumber terbuka berasaskan Python yang direka untuk membantu pembangun (*developer*) membangunkan aplikasi web dengan mudah dan pantas (Kochakadan, 2024). Salah satu kelebihan rangka kerja Streamlit ialah pengguna tidak perlu mengatur reka bentuk laman web menggunakan CSS, HTML, dan JavaScript kerana rangka kerja ini menyediakan fungsi yang diperlukan (Shukla et al., 2021).

Rajah 6 menunjukkan cara penggunaan aplikasi *ramalHujan*. Dalam langkah 5, graf tersebut memaparkan analisis data hujan menggunakan model SARIMA. Terdapat 2 komponen utama di papan pemuka (a) data hujan bermula dari tahun 1970 sehingga 2020 dan (b) unjuran ramalan hujan untuk tahun ke hadapan. Data hujan yang bermula dari tahun 1970–2014 diwakili oleh garis biru menunjukkan corak hujan bulanan yang berubah-ubah dengan pola bermusim yang jelas. Manakala

garis yang berwarna jingga pula mewakili unjuran ramalan corak hujan bulanan untuk tempoh masa ke hadapan selama 5 tahun (2015-2020) yang dibuat berdasarkan data sejarah hujan.



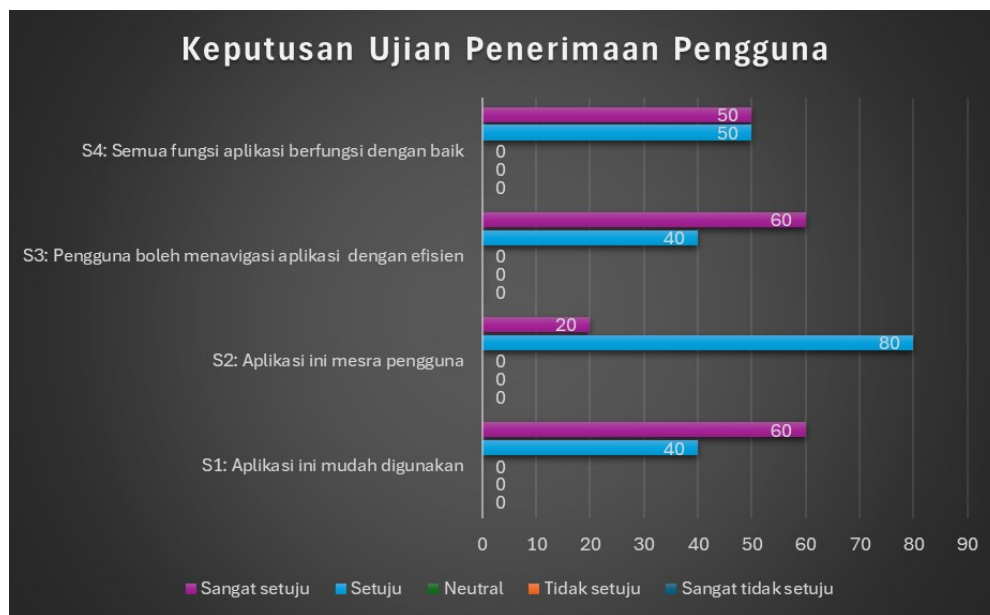
Rajah 6. Langkah atau tatacara menggunakan aplikasi *ramalHujan* bagi stesen Sungai Manggis, Selangor.

6. IMPLEMENTASI DAN PENGUJIAN

Rajah 7 memaparkan keputusan ujian penerimaan pengguna untuk aplikasi *ramalHujan* berdasarkan borang soal selidik telah dilengkapkan oleh 10 responden, termasuk pensyarah dan pelajar kursus TSV 3313 di Jabatan Sains Komputer. Berdasarkan dapatan kajian didapati: Soalan 1 (aplikasi ini mudah digunakan), majoriti responden (60%) sangat bersetuju bahawa aplikasi ini mudah digunakan. Manakala baki 40% menyatakan mereka setuju dengan kenyataan tersebut.

Soalan 2 (Aplikasi ini mesra pengguna), sebanyak 80% responden bersetuju aplikasi ini mesra pengguna, sementara 20% lagi sangat bersetuju. Soalan 3 (Pengguna boleh menavigasi aplikasi dengan efisien), di mana 60% responden sangat bersetuju bahawa aplikasi ini membolehkan navigasi yang pantas dan berkesan, manakala 40% lagi menyatakan persetujuan.

Soalan 4 (Semua fungsi aplikasi berfungsi dengan baik), separuh daripada responden (50%) sangat bersetuju, sementara baki 50% menyatakan persetujuan bahawa semua fungsi aplikasi beroperasi dengan baik. Secara keseluruhannya, maklum balas pengguna menunjukkan tahap kepuasan yang tinggi terhadap aplikasi *ramalHujan*. Majoriti responden menilai aplikasi ini sebagai mesra pengguna, mudah digunakan, dan berfungsi dengan efisien.



Rajah 7. Maklumbalas ujian penerimaan pengguna bagi aplikasi *ramalHujan*.

7. KESIMPULAN

Kajian ini berjaya membangunkan satu model ramalan hujan menggunakan model SARIMA berdasarkan data hujan bulanan dari stesen Sungai Manggis di Selangor. Model yang dibangunkan memberikan unjuran ramalan yang tepat serta dapat mengenal pasti pola bermusim yang konsisten dalam data hujan. Aplikasi *ramalHujan* yang diintegrasikan dengan model ini menyediakan kemudahan kepada pengguna untuk membuat ramalan hujan berdasarkan lokasi dan tempoh masa yang ditentukan. Bagi kajian masa hadapan, disarankan agar model sedia ada ini ditambahbaik dengan pengintegrasian data masa nyata (*real time data*) serta penambahan faktor lain seperti suhu, kelembapan, dan pola angin bagi mempertingkatkan ketepatan ramalan. Selain itu, aplikasi dan model ini juga berpotensi untuk digunakan di lokasi atau kawasan yang terdedah kepada risiko bencana alam seperti banjir. Pendekatan ini dapat memperkukuhkan peranan model dalam menyokong pembuatan keputusan berkaitan pengurusan sumber air dan strategi mitigasi bencana yang lebih berkesan.

Penghargaan:

Penulis mengucapkan penghargaan kepada Fakulti Sains dan Teknologi Pertahanan (FSTP) dan Universiti Pertahanan Nasional Malaysia di atas sokongan penyelidikan ini.

Rujukan

- Amin, M. N. M., Nohuddin, P. N., & Zainol, Z. (2014, December). Trend cluster analysis using self organizing maps. In *2014 4th World Congress on Information and Communication Technologies (WICT 2014)* (pp. 80-84). IEEE.
- Aworka, R., Cedric, L. S., Adoni, W. Y. H., Zoueu, J. T., Mutombo, F. K., Kimpolo, C. L. M., & Krichen, M. (2022). Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries. *Smart Agricultural Technology*, 2, 100048.
- Bassey, K. E. (2023). Hybrid renewable energy systems modeling. *Engineering Science & Technology Journal*, 4(6), 571-588.
- Dunham, M. H. (2006). Data mining: Introductory and advanced topics: Pearson Education India.
- Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques: Elsevier.
- Kalimuthu, M., Vaishnavi, P., & Kishore, M. (2020, August). Crop prediction using machine learning. In *2020 third international conference on smart systems and inventive technology (ICSSIT)* (pp. 926-932). IEEE.
- Kochakadan, A. (2024). Streamlit in Action Pure Python web apps in minutes: Manning Publication.
- Lai, Y., & Dzombak, D. A. (2020). Use of the autoregressive integrated moving average (ARIMA) model to forecast near-term regional temperature and precipitation. *Weather and forecasting*, 35(3), 959-976.
- Lim, S. E., Misiran, M., Sapiri, H., & Yusof, Z. M. (2024). Forecasting the Monthly Temperature and Rainfall in Chuping Perlis. *EDUCATUM Journal of Science, Mathematics and Technology*, 11(2), 20-33.
- Loh, W. S., Tan, W. L., Chin, R. J., Ling, L., Phoon, S. W., & Seah, C. S. (2024). An unsupervised machine learning approach for estimating missing daily rainfall data in peninsular malaysia. In *ITM Web of Conferences* (Vol. 67, p. 01010). EDP Sciences..
- Majka, M. (2024). Seasonal Time Series Analysis: Why SARIMA Outshines ARIMA.
- Marzukhi, S., Sidik, M. A. S. M., Nasir, H. M., Zainol, Z., & Ismail, M. N. (2018). Flood detection and warning system (FLoWS). In *Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication* (pp. 1-4).
- Mishra, A., & Desai, V. (2005). Drought forecasting using stochastic models. *Stochastic environmental research and risk assessment*, 19, 326-339.
- Mustapha, H., & Ismail, N. (2021). Time series modeling and forecasting of monthly rainfall using autoregressive integrated moving average (ARIMA) and seasonal autoregressive integrated moving average (SARIMA). *Proceedings of Science and Mathematics* (pp. 79–88).
- Mutalib, M. Y. A., Zainol, Z., Rauf, U. F. A., & Nohuddin, P. N. E. (2023). Analisis Ramalan Prestasi Akademik Pelajar Menggunakan Aplikasi MyCGPA. *Journal Information System and Technology Management (JISTM)*, 8(31), 80-94.
- Nohuddin, P. N., Zainol, Z., & Amir, M. A. U. (2018a). Trend cluster analysis of wave data for renewable energy. *Advanced Science Letters*, 24(2), 951-955.
- Nohuddin, P., Zainol, Z., Lee, A. S. H., Nordin, I., & Yusoff, Z. (2018). A case study in knowledge acquisition for logistic cargo distribution data mining framework. *International Journal of Advanced and Applied Sciences*, 5(1), 8-14.

- Nohuddin, P. N. E., Zainol, Z., & Nordin, A. (2018c). Monitoring students performance using self organizing map trend clustering. *Zulfaqr Journal of Defence Science, Engineering & Technology*, 1(1), 50-56.
- Noormanshah, W. M., Nohuddin, P. N., & Zainol, Z. (2021). Document content analysis based on random Forest algorithm. In *Intelligent Computing and Innovation on Data Science: Proceedings of ICTIDS 2019* (pp. 485-494). Singapore: Springer Singapore.
- Oddo, P. C., Bolten, J. D., Kumar, S. V., & Cleary, B. (2024). Deep Convolutional LSTM for improved flash flood prediction. *Frontiers in Water*, 6, 1346104.
- Poh, L. C., & Norrulashikin, S. M. (2022). Forecasting Amount of Rainfall in Peninsular Malaysia Using Holt-Winter Method and Seasonal Autoregressive Integrated Moving Average (SARIMA).
- Rahman, A. U., Abbas, S., Gollapalli, M., Ahmed, R., Aftab, S., Ahmad, M., Khan, M. A. & Mosavi, A. (2022). Rainfall prediction system using machine learning fusion for smart cities. *Sensors*, 22(9), 3504.
- Rashid, R. A., Nohuddin, P. N., & Zainol, Z. (2017). Association rule mining using time series data for Malaysia climate variability prediction. In *Advances in Visual Informatics: 5th International Visual Informatics Conference, IVIC 2017, Bangi, Malaysia, November 28–30, 2017, Proceedings 5* (pp. 120-130). Springer International Publishing.
- Rashid, R. A., Nohuddin, P. N., & Zainol, Z. (2018). Analyzing Climate Variability in Malaysia Using Association Rule Mining. *International Journal of Engineering & Technology*, 7(4.34), 394-397.
- Samsulaiman, N. H., & Ismail, N. (2023). Rainfall Modelling and Forecasting Using Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing in Pahang and Terengganu, Malaysia. *Proceedings of Science and Mathematics*, 18, 1-10.
- Shukla, S., Maheshwari, A., & Johri, P. (2021, December). Comparative analysis of ml algorithms & stream lit web application. In *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)* (pp. 175-180). IEEE.
- Tee, H. Y., & Mansor, R. (2024). Forecasting Rainfall Volume in Selangor with A Combined ARIMA Model. *Journal of Computational Innovation and Analytics (JCIA)*, 3(1), 83-103.
- Vun, K. J., Arumugasamy, S. K., Azari, M., & Yenn, T. F. (2024). Water Level Forecasting Using Artificial Neural Network (Ann): A Case Study of Semenyih River, Malaysia. *Process Integration and Optimization for Sustainability*, 8(1), 259-270.
- Wang, X., & Xie, H. (2018). A review on applications of remote sensing and geographic information systems (GIS) in water resources and flood risk management. *Water*, 10(5), 608.
- Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*. 1, 29-39.
- Yao, Z., Lum, Y., Johnston, A., Mejia-Mendoza, L. M., Zhou, X., Wen, Y., Guzik, A. A, Sargent, E. H., & Seh, Z. W. (2023). Machine learning for a sustainable energy future. *Nature Reviews Materials*, 8(3), 202-215.
- Zainol, Z., Jaymes, M. T., & Nohuddin, P. N. (2018). Visualurtext: a text analytics tool for unstructured textual data. In *Journal of Physics: Conference Series* (Vol. 1018, No. 1, p. 012011). IOP Publishing.
- Zainol, Z., Nohuddin, P. N., Kamarudin, N. D., Lee, A. S. H., & Wani, S. (2019). RELATIONSHIP ANALYSIS OF FORMAL AND EXPERIENTIAL LEARNING IN MILITARY SURVIVAL SKILLS USING TEXT MINING. *Defence S & T Technical Bulletin*, 12(2), 267-280.
- Zainudin, S., Jasim, D. S., & Bakar, A. A. (2016). Comparative analysis of data mining techniques for Malaysian rainfall prediction. *Int. J. Adv. Sci. Eng. Inf. Technol*, 6(6), 1148-1153.