

ANALISIS RAMALAN DAN PEMANTAUAN STATUS UNTUK AKTIVITI HARIAN MENGGUNAKAN PEMBELAJARAN MESIN (ML)

ARINA BINTI AZIZ

TS. DR. FADILLA 'ATYKA BINTI NOR RASHID

*Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia, 43600 UKM Bangi,
Selangor Darul Ehsan, Malaysia*

ABSTRAK

Kehidupan moden memerlukan pemantauan dan pengurusan aktiviti harian yang cekap untuk meningkatkan produktiviti dan mengurangkan kos. Projek ini membangunkan model pemantauan dan ramalan aktiviti harian menggunakan algoritma Long Short-Term Memory (LSTM) untuk menangani dua masalah utama iaitu kesukaran dalam mengesan dan menganalisis aktiviti secara sistematik serta cabaran dalam pemantauan berkesan oleh individu. Data dari dataset awam TensorFlow on Android for Human Activity Recognition digunakan, melibatkan sensor akselerometer untuk enam aktiviti utama. Platform Google Colab dipilih untuk pelaksanaan projek kerana kemudahan pengurusan kod dan sumber pengkomputeran. Model yang dibangunkan mencapai ketepatan ramalan 96%, menyediakan alat berkesan untuk memantau dan menganalisis aktiviti harian secara real-time, membantu pengguna membuat penyesuaian untuk meningkatkan gaya hidup dan produktiviti.

Kata kunci: analisis ramalan, pemantauan aktiviti, pembelajaran mesin, LSTM, Google Colab.

PENGENALAN

Pemantauan aktiviti harian adalah alat penting dalam era digital moden kerana ia membantu individu dan masyarakat menguruskan rutin harian mereka dengan lebih cekap. Dengan kemajuan teknologi, data sensor yang dikumpul secara berkala membolehkan ramalan yang tepat mengenai pola tingkah laku, membantu pengguna membuat keputusan yang lebih baik. Selain itu, pemantauan ini memainkan peranan penting dalam mengesan perubahan tingkah laku atau aktiviti, yang boleh menunjukkan masalah kesihatan atau keperluan penyesuaian gaya hidup. Dalam era digital yang pesat ini, pemantauan status bukan sahaja meningkatkan kesedaran tentang corak aktiviti harian tetapi juga menyokong produktiviti dan kesejahteraan keseluruhan.

Walau bagaimanapun, terdapat beberapa masalah utama yang perlu ditangani untuk meningkatkan pengurusan aktiviti harian. Kesukaran dalam mengesan dan menganalisis aktiviti secara sistematik menunjukkan keperluan untuk pendekatan berstruktur. Tambahan pula, cabaran dalam pemantauan berkesan menekankan keperluan sistem atau peranti yang lebih baik. Dengan gaya hidup sibuk yang sering mengabaikan aspek kesihatan, isu-isu ini menggalakkan pembangunan sistem yang lebih berkesan untuk memantau dan menganalisis data kesihatan secara menyeluruh.

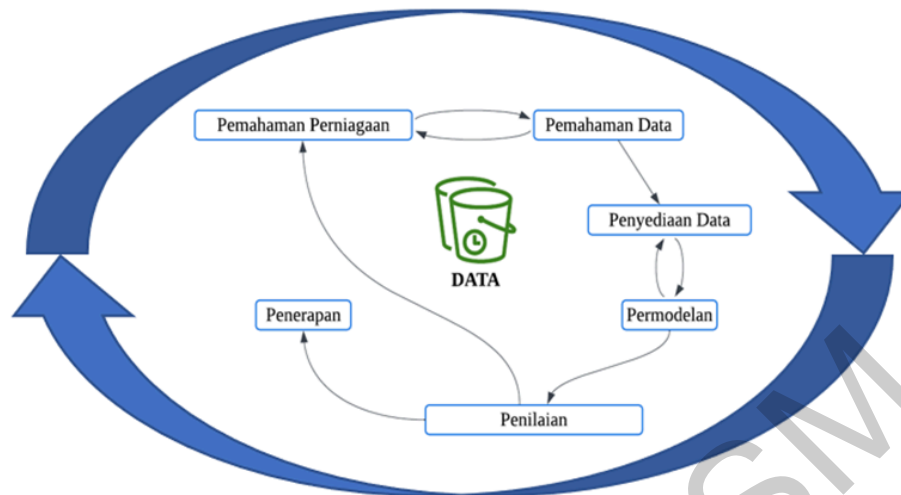
Kajian terdahulu menunjukkan bahawa penggunaan sensor untuk pemantauan aktiviti harian adalah penting untuk tujuan kesihatan, kecergasan, penjagaan kesihatan, dan keselamatan. Menurut Mahmud et al. (2020), sekitar 62% populasi global menggunakan peranti bersensor untuk pemantauan kesihatan, dan laporan Statista (2022) menunjukkan bahawa pasaran peranti boleh pakai dijangka mencapai 1.1 bilion unit menjelang 2024. Walaupun terdapat usaha dalam penggunaan sensor, masih ada keperluan untuk sistem yang lebih berkesan untuk mengatasi masalah yang dikenalpasti.

Sebagai langkah seterusnya, cadangan kerja ini melibatkan pembangunan sistem pemantauan aktiviti harian menggunakan teknologi sensor dan algoritma pembelajaran mesin. Sistem ini akan menggunakan data sensor yang dikumpul secara berkala untuk meramalkan dan menganalisis corak aktiviti harian. Inspirasi untuk menggunakan algoritma pembelajaran mesin, khususnya Long Short-Term Memory (LSTM), diperoleh dari kajian terdahulu yang menunjukkan keupayaan algoritma ini dalam menangani data bersiri dan memberikan ramalan yang tepat.

Secara keseluruhan, diharapkan hasil projek ini akan meningkatkan pemahaman tentang penggunaan sensor dalam kehidupan seharian dan menyumbang kepada perkembangan teknologi dalam aplikasi pemantauan aktiviti yang lebih canggih dan cekap. Keputusan akhir yang menunjukkan ketepatan tinggi dalam ramalan status aktiviti akan menyediakan alat yang berkesan untuk memantau dan menganalisis aktiviti harian secara real-time.

METODOLOGI KAJIAN

Metodologi CRISP-DM adalah model yang terkenal dan digunakan secara meluas dalam projek data mining dan pembelajaran mesin kerana fleksibiliti dan keupayaannya untuk menangani pelbagai jenis data dan masalah. Dengan struktur yang teratur dan pendekatan yang sistematik, CRISP-DM membolehkan analisis yang mendalam dan penyelesaian yang berkesan dalam pelbagai konteks. Menurut laporan oleh KDNuggets (2014), lebih daripada 43% projek data mining di seluruh dunia menggunakan CRISP-DM sebagai metodologi utama mereka, menjadikannya model yang paling banyak digunakan dalam industri. Model ini sesuai untuk menangani masalah dalam analisis ramalan dan pemantauan status dengan memberikan panduan yang jelas dalam setiap langkah proses data mining dan pembelajaran mesin.



Rajah menunjukkan fasa-fasa dalam CRISP-DM

Fasa Pemahaman Perniagaan (Business Understanding):

Pada fasa pertama ini, objektif utama adalah untuk mengenal pasti objektif dan keperluan perniagaan serta mentakrifkan masalah yang ingin diselesaikan. Projek ini bertujuan untuk meramalkan dan memantau aktiviti harian pengguna menggunakan data dari sensor akselerometer, dengan fokus pada tahap keaktifan seperti aktif, sederhana, dan tidak aktif. Ini penting bagi meningkatkan kualiti hidup dan produktiviti pengguna dengan menyediakan maklumat berguna tentang corak aktiviti harian mereka. Aktiviti yang akan dikenal pasti termasuk berjoging, berjalan, duduk, berdiri, memanjat tangga, dan menuruni tangga. Memahami tahap keaktifan ini adalah penting untuk memastikan penyelesaian yang dibangunkan dapat memenuhi tujuan perniagaan yang telah ditetapkan.

Fasa Pemahaman Data (Data Understanding):

Dalam fasa ini, objektifnya adalah untuk mengumpul data awal, mengenal pasti masalah dalam data, dan mendapatkan wawasan awal mengenainya. Data yang digunakan, yang dikenali sebagai "Sensor Data for Human Activity Recognition," diperoleh dari sensor akselerometer dan mengandungi 1,098,207 baris dan 6 lajur yang mencatatkan pelbagai

aktiviti fizikal. Fasa ini melibatkan pengumpulan dan penerokaan data untuk mengenal pasti corak, anomali, dan ciri-ciri utama yang boleh membantu dalam pemodelan aktiviti. Selain itu, data tersebut melibatkan 36 pengguna dan merakamkan jumlah pergerakan akselerometer untuk setiap jenis latihan fizikal, seperti berjoging (336445), berjalan (418393), duduk (59939), berdiri (48394), memanjat tangga (122869), dan menuruni tangga (100425). Dengan menganalisis data ini, kita dapat memahami lebih mendalam tentang struktur dan kandungan data, serta mengenal pasti elemen-elemen yang penting bagi menghasilkan model yang tepat untuk meramalkan dan memantau aktiviti harian pengguna.

Fasa Penyediaan Data (Data Preparation):

Dalam fasa ini, tumpuan adalah kepada membersihkan, mengintegrasikan, dan memformat data agar sesuai untuk proses pemodelan. Aktiviti utama dalam fasa ini termasuk pra-pemrosesan data dari sensor akselerometer, seperti menghapuskan outlier, menangani data yang hilang, dan menormalkan nilai pecutan untuk memastikan konsistensi dan kebolehpercayaan data. Sebagai contoh, teknik interpolasi digunakan untuk mengisi kekosongan dalam data dan penormalan dilakukan untuk memastikan data berada dalam skala yang seragam. Fasa ini bertujuan untuk menyediakan data yang bersih dan terstruktur, yang sangat penting untuk memastikan bahawa model yang dibangunkan adalah tepat dan boleh dipercayai dalam meramalkan serta memantau aktiviti harian pengguna.

Kod:

```
# Check for missing values before dropping
missing_values_before = df.isnull().sum()
print("Missing values before dropping:\n", missing_values_before)

# Drop NaN Values (removes any rows that contain missing values)
df = df.dropna()
```

Output Kod:

```
Missing values before dropping:  
user      0  
activity  0  
timestamp 0  
x-axis    0  
y-axis    0  
z-axis    1  
dtype: int64
```

Fasa Pemodelan (Modeling)

Fasa ini bertujuan untuk membangunkan dan menilai pelbagai model pembelajaran mesin untuk memilih yang paling sesuai. Dalam fasa ini, algoritma Long Short-Term Memory (LSTM) digunakan untuk membina model ramalan aktiviti harian berdasarkan data akselerometer. LSTM dipilih kerana kemampuannya dalam menangani data bersiri dan mengenal pasti corak jangka panjang dalam data sensor. Pada awal fasa, seni bina LSTM yang sesuai dibangunkan, terdiri daripada dua lapisan LSTM dengan 64 unit setiap satu dan dua lapisan fully-connected. Model ini dilatih menggunakan data latihan yang telah di pra-pemprosesan, dengan matlamat untuk mengenal pasti aktiviti seperti berjalan, berjoging, duduk, berdiri, naik tangga, dan turun tangga. Latihan dilakukan selama 50 epoch menggunakan TensorFlow, dan prestasi model dipantau melalui metrik ketepatan dan kehilangan (loss). Model LSTM menunjukkan prestasi yang memuaskan dengan ketepatan melebihi 97% dan kehilangan sekitar 0.2. Untuk meningkatkan prestasi, penyesuaian hyperparameter dilakukan, termasuk bilangan unit dalam setiap lapisan, bilangan epoch, dan kadar pembelajaran. Selepas latihan, model dinilai menggunakan set validasi dan set ujian untuk memastikan prestasi yang baik pada data yang tidak pernah dilihat semasa latihan, menghasilkan model yang berkesan untuk meramalkan dan memantau aktiviti harian pengguna.

Kod:

```

from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, LSTM, Dropout, Conv1D, MaxPooling1D, Flatten

#Sequential Model Initialization
model = Sequential()

#Convolutional Layers (CNN)
model.add(Conv1D(filters=64, kernel_size=3, activation='relu', input_shape=(time_steps, 3)))
model.add(Conv1D(filters=64, kernel_size=3, activation='relu'))
model.add(MaxPooling1D(pool_size=2))
model.add(Dropout(0.5))

#Long Short-Term Memory (LSTM)
model.add(LSTM(100))
model.add(Dropout(0.5))

#Dense layer
model.add(Dense(100, activation='relu'))
model.add(Dense(len(np.unique(y_encoded)), activation='softmax')) # Output layer

#Model Compilation
model.compile(loss='sparse_categorical_crossentropy', optimizer='adam', metrics=['accuracy'])

#Model Summary
print(model.summary())

```

```

history = model.fit(X_train, y_train, epochs=50, batch_size=1024,
                    validation_data=(X_test, y_test), verbose=1)

```

Output Kod:

```

Model: "sequential"
-----
Layer (type)                Output Shape              Param #
-----
conv1d (Conv1D)              (None, 98, 64)           640
conv1d_1 (Conv1D)            (None, 96, 64)           12352
max_pooling1d (MaxPooling1D) (None, 48, 64)           0
dropout (Dropout)            (None, 48, 64)           0
lstm (LSTM)                   (None, 100)              66000
dropout_1 (Dropout)          (None, 100)              0
dense (Dense)                 (None, 100)              10100
dense_1 (Dense)               (None, 6)                606
-----
Total params: 89698 (350.38 KB)
Trainable params: 89698 (350.38 KB)
Non-trainable params: 0 (0.00 Byte)
-----
None

```

```

Epoch 42/50
15/15 [=====] - 17s 1s/step - loss: 0.1155 - accuracy: 0.9615 - val_loss: 0.1089 - val_accuracy: 0.9669
Epoch 43/50
15/15 [=====] - 14s 968ms/step - loss: 0.1206 - accuracy: 0.9586 - val_loss: 0.1153 - val_accuracy: 0.96
Epoch 44/50
15/15 [=====] - 14s 966ms/step - loss: 0.1109 - accuracy: 0.9629 - val_loss: 0.1022 - val_accuracy: 0.96
Epoch 45/50
15/15 [=====] - 14s 966ms/step - loss: 0.1071 - accuracy: 0.9652 - val_loss: 0.1032 - val_accuracy: 0.96
Epoch 46/50
15/15 [=====] - 15s 1s/step - loss: 0.1081 - accuracy: 0.9641 - val_loss: 0.1111 - val_accuracy: 0.9646
Epoch 47/50
15/15 [=====] - 16s 1s/step - loss: 0.1096 - accuracy: 0.9642 - val_loss: 0.1196 - val_accuracy: 0.9623
Epoch 48/50
15/15 [=====] - 16s 1s/step - loss: 0.1048 - accuracy: 0.9659 - val_loss: 0.1014 - val_accuracy: 0.9690
Epoch 49/50
15/15 [=====] - 15s 1s/step - loss: 0.1067 - accuracy: 0.9642 - val_loss: 0.0957 - val_accuracy: 0.9712
Epoch 50/50
15/15 [=====] - 15s 1s/step - loss: 0.0982 - accuracy: 0.9674 - val_loss: 0.1167 - val_accuracy: 0.9640

```

Fasa Penilaian (Evaluation)

Seterusnya, fasa ini menumpukan kepada penilaian model yang telah dibangunkan untuk memastikan ia memenuhi objektif perniagaan dan berfungsi dengan baik. Dalam konteks ini, model LSTM yang dibangunkan untuk meramalkan aktiviti harian dinilai menggunakan pelbagai metrik prestasi, termasuk ketepatan, ketelitian, dan skor F1. Model ini diuji secara menyeluruh pada set ujian yang tidak pernah dilihat semasa latihan untuk mengesahkan keupayaannya dalam memberikan ramalan yang tepat. Penilaian ini bertujuan untuk memastikan bahawa model tidak hanya berfungsi dengan baik dalam persekitaran latihan tetapi juga dapat digunakan secara efektif dalam aplikasi sebenar, memenuhi keperluan perniagaan yang telah ditetapkan.

Kod:

```

#Classification Report
print(classification_report(y_test, y_pred))

#Confusion Matrix
cm = confusion_matrix(y_test, y_pred)
plt.figure(figsize=(10, 7))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=encoder.classes_,
            yticklabels=encoder.classes_)
plt.title('Confusion Matrix')
plt.xlabel('Predicted Labels')
plt.ylabel('True Labels')
plt.show()

```


Output Kod:

	precision	recall	f1-score	support
0	0.83	0.95	0.89	623
1	1.00	0.98	0.99	2032
2	0.98	0.99	0.98	343
3	0.98	0.96	0.97	302
4	0.92	0.84	0.87	714
5	0.98	0.99	0.99	2505
accuracy			0.96	6519
macro avg	0.95	0.95	0.95	6519
weighted avg	0.97	0.96	0.96	6519

Fasa Penggunaan (Deployment):

Akhir sekali, fasa ini menumpukan kepada penerapan model dalam persekitaran sebenar dan pemantauan prestasinya. Dalam projek ini, model LSTM yang telah berjaya dibangunkan akan diintegrasikan ke dalam aplikasi atau platform yang membolehkan pengguna mengakses ramalan dan pemantauan aktiviti harian mereka. Sebagai contoh, aplikasi mudah alih atau laman web akan disediakan untuk membolehkan pengguna memantau aktiviti harian mereka secara real-time berdasarkan ramalan model.

Data yang diperoleh daripada penilaian kebolegunaan akan dianalisis dengan menggunakan kaedah statistik deskriptif. Penilaian ini melibatkan pembahagian prestasi kepada kategori seperti:

- Sangat Baik: > 400,000
- Baik: 250,000 - 400,000
- Memuaskan: 100,000 - 250,000
- Sederhana: < 100,000

Jadual berikut menunjukkan tafsiran skala skor setiap aktiviti untuk keseluruhan pengguna

Aktiviti	Skor Min	Tafsiran
Berjalan	418,393	Sangat Baik
Berjoging	336,445	Baik
Duduk	59,939	Sederhana
Berdiri	48,394	Sederhana
Naik Tangga	122,869	Memuaskan
Turun Tangga	100,425	Memuaskan

Fasa ini memastikan model dapat digunakan secara berkesan dalam kehidupan harian pengguna, memberikan alat yang berguna untuk memantau dan menganalisis aktiviti dengan ketepatan yang tinggi.

Secara keseluruhannya, CRISP-DM menekankan kepentingan memahami objektif perniagaan dan memastikan hasil akhir projek selaras dengan matlamat perniagaan. Seperti yang dinyatakan oleh Chapman et al. (2000), CRISP-DM adalah metodologi yang paling banyak digunakan dalam projek data mining, dengan peratusan penggunaan melebihi 50% dalam pelbagai industri. Dalam projek ini, CRISP-DM telah dipilih kerana strukturnya yang sistematik, fleksibiliti, dan penerimaan luas dalam industri. Hasil akhir projek ini adalah model pembelajaran mesin yang mampu meramalkan dan memantau aktiviti harian dengan ketepatan tinggi. Model ini akan diintegrasikan ke dalam aplikasi pemantauan aktiviti, memberikan alat yang berguna untuk membantu pengguna memahami dan mengurus aktiviti harian mereka dengan lebih baik. Dengan demikian, projek ini diharapkan dapat menyumbang secara signifikan kepada perkembangan teknologi dalam aplikasi pemantauan aktiviti yang lebih canggih dan cekap.

KEPUTUSAN DAN PERBINCANGAN

Visualisasi data ini dirancang untuk meningkatkan kecekapan pengguna dalam memantau dan menganalisis aktiviti harian mereka. Dengan memanfaatkan data dari sensor akselerometer, visualisasi ini memudahkan ramalan dan pemantauan aktiviti seperti berjalan, berjoging, menaiki tangga, menuruni tangga, duduk, dan berdiri. Dibangunkan menggunakan Google Colab, Python, dan pustaka TensorFlow untuk model LSTM, platform ini mengendalikan pemodelan dan analisis data. Data akselerometer, yang dikumpul dari sensor pada peranti pengguna, dianalisis untuk mengenal pasti corak aktiviti. Aliran keputusan dan perbincangan antara muka ini merangkumi analisis corak dan maklumat berguna yang diberikan kepada pengguna tentang aktiviti harian mereka, memastikan sistem ini memberikan maklumat yang tepat dan relevan.

Pengguna dapat memilih nama mereka dari senarai dropdown yang memaparkan semua nama pengguna yang terdapat dalam sistem, dari User 1 hingga User 36. Setelah memilih nama, senarai hari yang tersedia akan dikemaskini secara automatik berdasarkan data yang ada dalam pangkalan data. Pengguna kemudiannya boleh memilih hari tertentu dari senarai dropdown yang dikemaskini, yang terdiri daripada Khamis, Jumaat, dan Sabtu, untuk melihat aktiviti harian mereka dengan lebih terperinci. Proses ini memastikan hanya data yang relevan dipaparkan kepada pengguna, memudahkan navigasi dan analisis aktiviti harian mereka dengan lebih efisien.



The image shows a dark-themed user interface with two dropdown menus. The first menu is labeled 'User:' and has 'User 21' selected. The second menu is labeled 'Day:' and has 'Friday' selected. Both menus have a small downward-pointing arrow on the right side.

Rajah menunjukkan pemilihan Pengguna 21 dan Hari Khamis

Setelah pengguna memilih nama dan hari yang relevan, sistem akan mengira dan menentukan status keseluruhan aktiviti pengguna berdasarkan data yang dikumpulkan sepanjang hari tersebut. Setiap aktiviti harian akan dinilai dan dikategorikan sebagai Active, Moderate, atau Inactive mengikut julat nilai yang telah ditetapkan.

Aktiviti	Inactive	Moderate	Active
Walking	0 - 4,999	5,000 - 9,999	>10,000
Jogging	0 - 2,999	3,000 - 5,999	>6,000
Sitting	0 - 999	1,000 - 1,999	>2,000
Standing	0 - 999	1,000 - 1,999	>2,000
Menaiki Tangga	0 - 1,999	2,000 - 3,999	>4,000
Menuruni Tangga	0 - 1,999	2,000 - 3,999	>4,000

Penilaian ini membolehkan pengguna mendapatkan pandangan yang jelas mengenai tahap aktiviti mereka pada hari tersebut, membantu mereka dalam memantau dan menilai corak aktiviti harian mereka dengan lebih baik.

Secara keseluruhan, sistem ini menganalisis jumlah skor yang diperoleh dari setiap jenis aktiviti iaitu Aktif, Sederhana, dan Tidak Aktif. Skor keseluruhan kemudiannya dikategorikan kepada tiga tahap utama:

- Aktif: Jika jumlah skor melebihi atau sama dengan 32, menandakan bahawa pengguna telah aktif sepanjang hari.
- Sederhana: Jika jumlah skor berada antara 24 hingga 31, menunjukkan tahap aktiviti yang sederhana.
- Tidak Aktif: Jika jumlah skor kurang daripada 24, menunjukkan bahawa pengguna kurang aktif sepanjang hari.

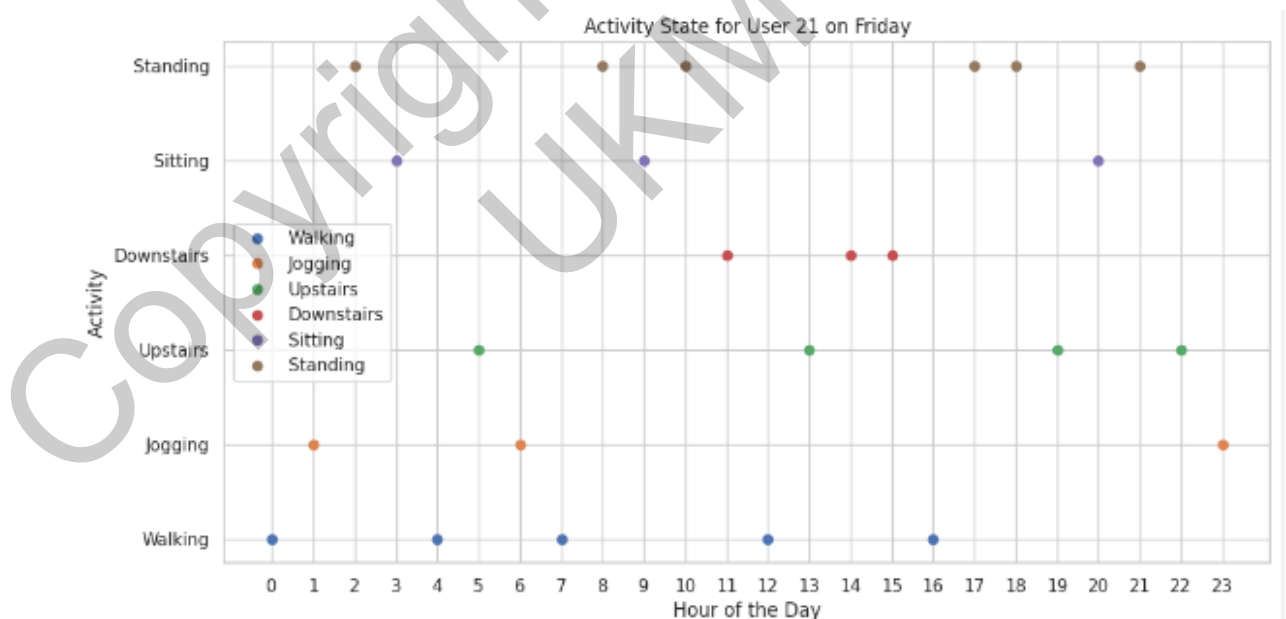
Status keseluruhan ini dipaparkan kepada pengguna untuk memberikan gambaran ringkas tentang tahap aktiviti mereka pada hari tersebut, membolehkan mereka memahami corak aktiviti harian mereka dengan lebih baik dan mengambil langkah yang sesuai untuk meningkatkan kualiti hidup mereka.

Overall status for User 21 on Friday: Active

Rajah Visualisasi Data Penentuan Status Keseluruhan bagi Pengguna 21 pada Hari Khamis

Untuk memperjelas analisis aktiviti, visualisasi data ini menyediakan dua jenis visualisasi yang membantu pengguna memahami corak aktiviti mereka dengan lebih mendalam.

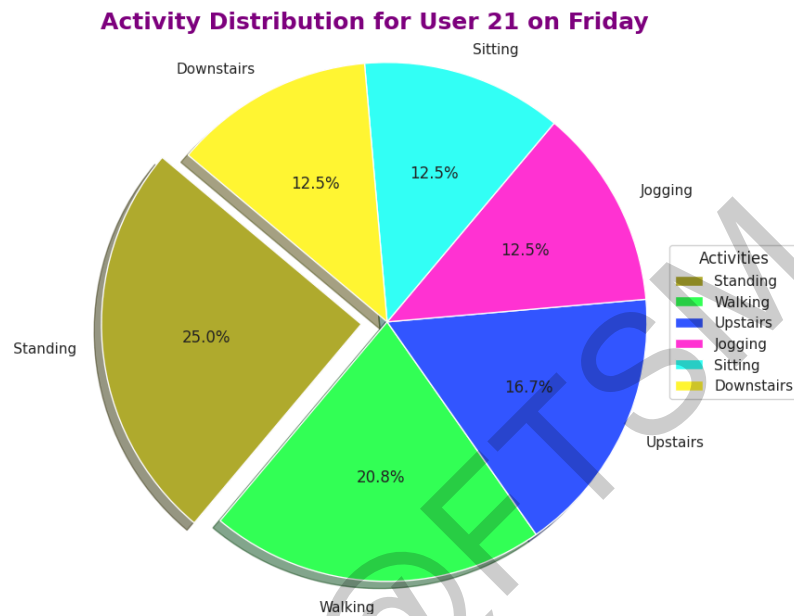
- Graf Aktiviti: Memaparkan aktiviti pengguna setiap jam sepanjang hari dengan warna dan simbol yang berbeza. Ini membolehkan pengguna melihat corak aktiviti mereka secara terperinci dan menganalisis perubahan aktiviti sepanjang hari.



Rajah Menunjukkan Graf Aktiviti seharian untuk pengguna 21 pada hari Jumaat

- Carta Pai Aktiviti: Peratusan masa yang dibahagikan antara pelbagai aktiviti seperti berjalan, berjoging, dan duduk. Ini memberikan gambaran ringkas tentang

pembahagian masa, membantu pengguna membuat keputusan mengenai perubahan gaya hidup yang diperlukan.



Rajah Menunjukkan Carta Pai Aktiviti seharian untuk pengguna 21 pada hari Jumaat

Visualisasi data ini direka untuk memberikan pengalaman yang responsif dan sentiasa terkini. Setiap kali pengguna memilih hari baru atau membuat perubahan pilihan, sistem secara automatik mengemaskini semua elemen antara muka, termasuk graf aktiviti dan carta pai, untuk memaparkan data yang relevan tanpa memerlukan tindakan tambahan. Kemas kini dinamik ini memastikan pengguna sentiasa mendapat maklumat yang tepat dan terkini mengenai aktiviti harian mereka. Dengan menyediakan visualisasi yang jelas dan maklumat yang mudah difahami, antara muka ini berfungsi sebagai alat yang berkuasa untuk memantau dan menganalisis aktiviti, membantu pengguna membuat keputusan yang lebih baik mengenai gaya hidup dan tahap aktiviti mereka. Melalui pemantauan berterusan dan penyesuaian yang diperlukan, pengguna dapat mencapai matlamat kesihatan dan kecergasan mereka dengan lebih berkesan.

Pengujian Kebolegunaan

Pengujian kebolegunaan sistem ini melibatkan beberapa langkah kritikal untuk memastikan bahawa antara muka dan fungsi sistem memenuhi keperluan pengguna dengan efektif. Pertama, ujian antaramuka pengguna dilakukan untuk menilai kemudahan penggunaan widget interaktif seperti dropdown bagi pemilihan pengguna dan hari, serta kejelasan maklumat yang dipaparkan. Kriteria utama dalam ujian ini termasuk kesederhanaan navigasi, ketepatan maklumat, dan responsif antaramuka. Kedua, ujian ketepatan data dilaksanakan untuk memastikan bahawa pengiraan aktiviti dan status keseluruhan berdasarkan data sensor adalah tepat, dengan membandingkan hasil output sistem dengan data yang dijana secara manual dan memastikan graf serta carta pai memaparkan gambar yang akurat tentang aktiviti pengguna. Ketiga, ujian fungsi memastikan semua ciri sistem berfungsi seperti yang diharapkan, termasuk kemaskini dinamik pada pilihan hari berdasarkan pengguna yang dipilih dan paparan graf serta carta yang relevan. Pengujian dilakukan menggunakan pelbagai set data, dengan jadual nilai yang dihasilkan menunjukkan perubahan dalam pengiraan aktiviti dan status.

Aktiviti dan Status Pengguna pada Hari Khamis

Pengguna	Jam Aktiviti						Jumlah Skor	Status Keseluruhan
	Berjalan	Jogging	Duduk	Berdiri	Menaiki Tangga	Menuruni Tangga		
1	4500	3200	1200	800	2500	1500	12,200	Moderate
2	5500	3500	1500	1000	3000	2000	17,000	Active
3	6000	4000	1300	900	2800	1700	20,700	Active
4	5000	3000	1500	1100	2400	1600	15,600	Moderate
5	4800	3200	1400	850	2600	1800	15,650	Moderate
6	5800	3600	1600	900	2900	1900	18,800	Active
7	4000	2500	900	700	2000	1300	10,400	Inactive

8	4500	3700	1300	800	2700	1600	14,600	Moderate
9	5500	3900	1400	950	2800	1700	18,250	Active
10	5700	3800	1500	1000	3000	1600	19,600	Active
11	6000	4000	1400	1100	3100	1800	22,400	Active
12	4600	3400	1300	850	2500	1500	13,750	Moderate
13	4700	3500	1200	950	2400	1600	14,350	Moderate
14	5900	3700	1400	1100	2800	1900	20,000	Active
15	4800	3200	1300	900	2300	1600	14,100	Moderate
16	5500	3600	1400	1000	2900	1800	18,200	Active
17	5600	3800	1300	1100	2700	1900	19,400	Active
18	4700	3500	1200	950	2400	1700	14,450	Moderate
19	4800	3400	1300	1000	2500	1600	14,600	Moderate
20	4600	3200	1400	950	2400	1500	13,500	Moderate
21	5500	3600	1500	1000	2700	1700	16,000	Active
22	3800	3000	900	800	2200	1400	9100	Inactive
23	4700	3400	1300	950	2500	1600	14,450	Moderate
24	4600	3300	1400	900	2600	1500	13,800	Moderate
25	4700	3400	1300	950	2500	1600	14,450	Moderate
26	4900	3200	1300	1000	2400	1500	14,300	Moderate
27	4700	3300	1300	950	2600	1600	14,450	Moderate
28	4800	3400	1400	950	2500	1500	14,550	Moderate
29	4600	3200	1300	1000	2400	1500	13,000	Moderate
30	4700	3300	1300	950	2400	1500	13,150	Moderate
31	5000	3100	1200	900	2300	1600	13,100	Moderate
32	4700	3400	1300	950	2400	1600	13,800	Moderate
33	4900	3300	1300	1000	2500	1500	13,500	Moderate
34	4600	3100	1200	900	2200	1400	12,500	Moderate
35	4700	3200	1300	950	2300	1600	13,050	Moderate
36	4800	3400	1400	950	2400	1500	14,450	Moderate

Cadangan Penambahbaikan

Untuk meningkatkan sistem analisis dan pemantauan aktiviti harian menggunakan pembelajaran mesin, beberapa cadangan penambahbaikan yang lebih spesifik boleh dipertimbangkan. Pertama, pengumpulan data yang lebih luas dari pelbagai sumber, seperti sensor yang berbeza dan dalam pelbagai situasi persekitaran, dapat meningkatkan ketepatan ramalan dengan menyediakan konteks yang lebih menyeluruh. Kedua, penerapan algoritma pembelajaran mesin yang lebih efisien, seperti yang mengurangkan keperluan pengkomputeran atau yang lebih sesuai untuk pemrosesan masa nyata, dapat mengoptimalkan penggunaan sumber dan mempercepat prestasi sistem. Ketiga, pengembangan teknik skalabiliti, seperti penggunaan seni bina mikro servis, akan membolehkan sistem menangani peningkatan bilangan pengguna tanpa menjejaskan prestasi. Selain itu, integrasi dengan teknologi Internet of Things (IoT) untuk mendapatkan data secara real-time akan memperkayakan analisis dan memungkinkan respons yang lebih cepat dan relevan. Akhirnya, melibatkan pengguna secara aktif dalam proses penambahbaikan antara muka, melalui maklum balas berterusan dan ujian kegunaan, akan memastikan sistem ini lebih mesra pengguna dan intuitif. Dengan melaksanakan cadangan-cadangan ini, sistem dapat mencapai tahap prestasi yang lebih tinggi dan memberikan manfaat yang lebih besar kepada penggunanya.

KESIMPULAN

Secara keseluruhannya, projek ini telah berjaya membangunkan sistem analisis ramalan dan pemantauan status aktiviti harian menggunakan pembelajaran mesin dengan aplikasi algoritma CNN-LSTM. Sistem ini telah membuktikan kemampuannya dalam meramalkan dan memantau aktiviti harian pengguna dengan ketepatan yang lebih tinggi. Proses pembangunan telah dilaksanakan dengan teliti, bermula dari pengekodan berdasarkan

spesifikasi yang telah ditetapkan, hingga reka bentuk yang diusulkan. Fasa pengujian juga telah dijalankan dengan menyeluruh, termasuk ujian keselamatan, prestasi, dan kesesuaian fungsi, untuk memastikan kualiti dan kesahihan sistem. Hasil daripada usaha ini, sistem yang dibangunkan bukan sahaja berfungsi dengan baik tetapi juga sesuai untuk digunakan dalam situasi sebenar, memenuhi matlamat utama projek ini dan menyediakan nilai tambah yang signifikan kepada pengguna dalam memantau dan meningkatkan gaya hidup harian mereka.

Kekuatan Sistem

Sistem ini menonjol dengan ketepatan ramalan yang tinggi, memanfaatkan algoritma CNN-LSTM untuk meramalkan aktiviti harian pengguna dengan tepat, memberikan ramalan yang relevan dan dapat diandalkan. Keupayaan pemantauan masa nyata membolehkan pengesanan perubahan dalam corak aktiviti secara langsung, menyediakan maklumat terkini yang sangat berguna, terutama dalam situasi yang memerlukan tindakan segera. Antara muka pengguna yang mesra dan inisiatif memudahkan interaksi, menjadikannya mudah digunakan oleh pelbagai golongan tanpa memerlukan pengetahuan teknikal yang mendalam. Selain itu, sistem ini dilengkapi dengan mekanisme pengesanan ralat awal yang berkesan, hasil daripada pembangunan terperinci dan pengujian menyeluruh, memastikan kestabilan dan kelancaran sistem serta meningkatkan pengalaman pengguna secara keseluruhan.

Kelemahan Sistem

Berikut adalah beberapa kelemahan dalam pelaksanaan sistem analisis dan pemantauan aktiviti harian menggunakan algoritma CNN-LSTM. Pertama, keberkesanan sistem bergantung kepada kualiti dan kuantiti data yang tersedia untuk melatih model; kekurangan data atau data berkualiti rendah boleh merendahkan prestasi sistem. Kedua, algoritma

CNN-LSTM memerlukan sumber pengkomputeran yang tinggi, yang mungkin menjadi halangan bagi pengguna yang tidak mempunyai akses kepada infrastruktur komputer yang memadai. Ketiga, sistem mungkin menghadapi cabaran dalam menguruskan skala besar, terutama apabila digunakan oleh banyak pengguna secara serentak, yang dapat mempengaruhi kehandalan dan prestasi. Memahami kelemahan-kelemahan ini penting untuk memperbaiki dan meningkatkan sistem dengan strategi yang tepat.

PENGHARGAAN

Syukur kepada Allah SWT di atas kesihatan dan kekuatan yang diberikan dalam menyiapkan projek ini. Begitu juga, kepada penyelia saya, Ts. Dr. Fadilla 'Atyka Binti Nor Rashidatas bantuan yang begitu besar, teguran dan nasihat yang begitu berguna sepanjang kajian ini.

Terima kasih juga saya ucapkan kepada pihak fakulti, para pensyarah yang telah bertungkus lumus mengajar saya pelbagai ilmu yang berguna dalam menyelesaikan projek ini.

Tidak lupa juga kepada keluarga dan sahabat yang tidak putus memberikan semangat dan bimbingan sepanjang penghasilan projek ini.

RUJUKAN

Mahmud, S., Kaiser, M. S., Hussain, A., & Vasilakos, A. V. (2020). "Wearable sensors for remote health monitoring." *Sensors*, 20(18), 5234.

Statista. (2022). "Wearable device unit sales worldwide from 2014 to 2024 (in millions)." Statista.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). "CRISP-DM 1.0: Step-by-step data mining guide." SPSS Inc.

Provost, F., & Fawcett, T. (2013). "Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking." O'Reilly Media.

Kotsiantis, S. B., Kanellopoulos, D., & Pintelas, P. E. (2006). "Data preprocessing for supervised learning." *International Journal of Computer Science*.

Hochreiter, S., & Schmidhuber, J. (1997). "Long Short-Term Memory." *Neural Computation*, 9(8), 1735-1780.

Kohavi, R., & Provost, F. (1998). "Glossary of terms." *Machine Learning*, 30(2-3), 271-274.

Shmueli, G., Patel, N. R., & Bruce, P. C. (2010). "Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner." John Wiley & Sons.

Mahmud, S., Kaiser, M. S., Hussain, A., & Vasilakos, A. V. (2020). "Wearable sensors for remote health monitoring." *Sensors*, 20(18), 5234.

Statista. (2022). "Wearable device unit sales worldwide from 2014 to 2024 (in millions)." Statista.

Arina Binti Aziz (A186664)

Ts. Dr. Fadilla 'Atyka Binti Nor Rashid

Fakulti Teknologi & Sains Maklumat

Universiti Kebangsaan Malaysia