

# RAMALAN SIRI MASA HARGA MATA WANG BITCOIN DENGAN MENGUNAKAN PENDEKATAN PEMBELAJARAN MESIN

Eddie Ngai & Salwani Abdullah

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

## Abstrak

Sejak beberapa tahun kebelakangan ini, Bitcoin telah menarik perhatian pelbagai pihak, daripada penyelidik akademik hinggalah kepada pelabur institusi. Bitcoin ialah mata wang kripto yang pertama dan paling banyak digunakan setakat ini. Disebabkan oleh turun naik yang ketara pada harga Bitcoin dan hakikat bahawa kaedah dagangannya tidak memerlukan pihak ketiga, ia telah mendapat banyak populariti sejak penubuhannya pada tahun 2009 di kalangan pelbagai individu. Memandangkan kesukaran sebelum ini dalam meramalkan harga mata wang kripto, projek ini akan membangun dan melaksanakan ramalan siri masa penyelesaian berasaskan pendekatan pembelajaran mesin yang merangkumi Sokongan Vektor Mesin Regresi (SVR), K-Jiran Terdekat Regresi (KNN), Peningkatan Kecerunan Melampau (XGBoost), dan Memori Jangka Pendek Panjang (LSTM) untuk menentukan arah aliran pergerakan harga Bitcoin dan menilai keberkesanan model pembelajaran Bitcoin. Data yang akan digunakan ialah harga penutupan Bitcoin dari tahun 2018 sehingga tahun 2023. Prestasi model pembelajaran mesin akan dinilai dengan membandingkan keputusan R kuasa dua, min ralat mutlak (MAE), punca ralat kuasa dua min (RMSE), dan juga melalui graf siri masa visualisasi harga penutupan asal dan ramalan harga tutup Bitcoin dalam papan pemuka. Antara model yang dibandingkan, LSTM muncul sebagai yang paling tepat, diikuti oleh SVR, manakala XGBoost dan KNN mempamerkan prestasi yang agak rendah.

## Pengenalan

Peningkatan mendadak Bitcoin telah memikat kedua-dua pelabur dan penyelidik. Walau bagaimanapun, ketidaktentuan harga yang melampau telah menimbulkan cabaran yang besar untuk ramalan yang tepat, menjadikannya masalah penting untuk diselesaikan. Dengan turun naik Bitcoin yang tinggi dan populariti yang semakin meningkat sebagai aset pelaburan, model ramalan yang boleh dipercayai boleh membantu pelabur dalam membuat keputusan termaklum, menyumbang kepada kecekapan pasaran dan memajukan pemahaman kita tentang dinamik mata wang kripto. Penemuan ini boleh diaplikasikan dengan meluas dalam kewangan, dan memberi inspirasi kepada penyelidikan lanjut.

Matlamat utama projek ini adalah untuk membina model pembelajaran mesin yang boleh dipercayai dan tepat yang meramalkan harga Bitcoin masa hadapan berdasarkan data sejarah Bitcoin terkini. Bagi mencapai matlamat ini, beberapa objektif telah disenaraikan. Objektif kajian ini adalah untuk menilai kesan nisbah pembahagian data latih-uji yang berbeza dan penalaan hiperparameter terhadap keputusan model ramalan, lalu membandingkan prestasi dan keputusan model pembelajaran mesin yang berbeza pada ramalan siri masa Bitcoin. Model ramalan tersebut juga diuji kebolehgunaannya dengan menggunakan set data mata wang kripto yang berbeza.

Walaupun kajian ini memberi tumpuan khusus pada ramalan harga Bitcoin siri masa, adalah penting untuk mengakui batasan yang wujud dalam projek itu. Skop adalah terhad kepada data sejarah harga harian Bitcoin selama 5 tahun dan tidak menggabungkan faktor luaran yang mungkin mempengaruhi pasaran mata wang kripto, seperti perubahan peraturan atau peristiwa ekonomi global. Walaupun terdapat batasan ini, penyelidikan ini bertujuan untuk mewujudkan asas yang kukuh untuk membangunkan model ramalan dalam domain mata wang kripto.

Kepentingan projek ini terletak pada potensi kesannya terhadap kedua-dua komuniti saintifik dan kewangan. Dengan berjaya membina model ramalan harga Bitcoin siri masa yang tepat, pengkaji boleh meningkatkan pemahaman tentang corak asas yang mengawal pergerakan harga mata wang

kripto. Selain itu, penemuan kajian ini boleh membantu pelabur dalam mengurangkan risiko dan memaksimumkan keuntungan dalam pasaran mata wang kripto yang sentiasa berkembang.

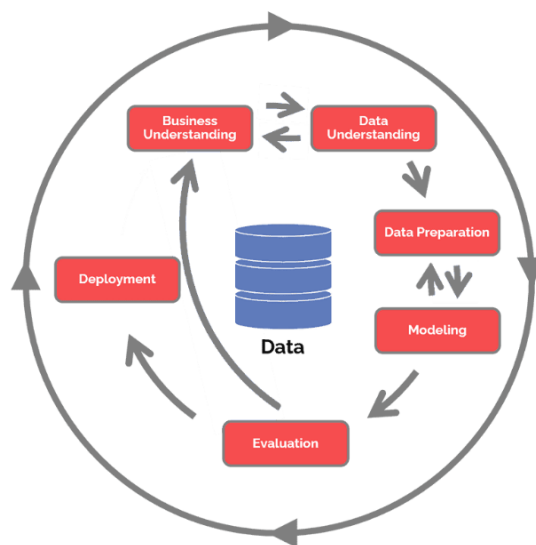
Proses pembangunan untuk projek ini akan mengikut metodologi CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Pendekatan yang diiktiraf secara meluas dan berstruktur ini akan membimbing peringkat penerokaan, penyediaan data, pembinaan model, penilaian dan penggunaan pengkaji. Dengan mematuhi rangka kerja yang ketat ini, pengkaji boleh memastikan kebolehulungan dan kebolehpercayaan hasil.

Terdapat beberapa kajian lepas yang telah dirujuk untuk mendapat gambaran tentang kajian dilakukan. Kajian pertama bertajuk Analisis Siri Masa Harga Mata Wang Kripto Menggunakan Memori Jangka Pendek Panjang (LSTM) yang diterbitkan pada 2022 (Fleischer et al., 2022), kajian ini menunjukkan LSTM menghasilkan RMSE yang lebih rendah (1334.755) untuk set data Bitcoin daripada model lain dengan kos masa jalan yang lebih lama. Kajian seterusnya bertajuk Ramalan Harga Mata Wang Kripto: Perbandingan Pembelajaran Mesin, Pembelajaran Mendalam dan Ensemble (Murray et al., 2023), kajian ini menunjukkan pembelajaran mendalam seperti LSTM mempamerkan RMSE 0.02224 dan Skor R<sup>2</sup> 0.735, mengatasi KNN, RMSE sebanyak 0.02332 dan Skor R<sup>2</sup> 0.711, serta Regresi Vektor Sokongan (SVR) RMSE 0.02452 dan Skor R<sup>2</sup> 0.681. Kajian ketiga bertajuk Ramalan siri masa harga Bitcoin menggunakan ciri dimensi tinggi: pendekatan pembelajaran mesin yang diterbitkan pada 2020 (Mudassir et al., 2020), kajian ini menunjukkan prestasi tinggi model klasifikasi dan regresi berasaskan pembelajaran mesin untuk meramalkan pergerakan harga dan harga Bitcoin dimana model Rangkaian Saraf Tiruan Bertindan (SANN) melaporkan ralat RMSE terendah (140.00) untuk ufuk ramalan 90 hari, diikuti oleh SVM(203.11), Rangkaian Saraf Tiruan (ANN) (210.09) dan LSTM (217.84). Penyelidikan seterusnya yang dikaji adalah bertajuk Meramalkan Harga Bitcoin Menggunakan Pembelajaran Mesin yang diterbitkan pada 2018 (McNally et al., 2018), menunjukkan LSTM mengatasi prestasi Rangkaian Saraf Berulang (RNN), tetapi tidak ketara dari segi ketepatan, di mana ia adalah 52.78% dan 50.25% masing-masing.

Laporan teknikal mengikut organisasi berstruktur, yang terdiri daripada empat bahagian utama. Pertama, Pengenalan memberikan gambaran keseluruhan latar belakang projek, objektif, skop, justifikasi, dan metodologi yang diguna pakai, menggariskan kepentingan ramalan harga Bitcoin siri masa. Bahagian Metodologi menerangkan model proses pembangunan yang digunakan, kaedah pengumpulan data dan analisis data serta pengukuran dan alat ukur keputusan kajian. Dalam bahagian Keputusan dan Perbincangan, prestasi model pembelajaran mesin dinilai secara menyeluruh, melalui penilaian metrik R kuasa dua ( $R^2$ ), min ralat mutlak (MAE), punca ralat kuasa dua min (RMSE), dan juga melalui graf siri masa visualisasi. Akhir sekali, kesimpulannya meringkaskan penemuan utama, implikasi terhadap industri, dan cadangan untuk kajian masa hadapan yang boleh dilakukan.

### Metodologi Kajian

Bagi metodologi projek ini, model perlombongan data yang boleh dipercayai akan dilaksanakan iaitu Proses Standard Merentas Industri untuk Perlombongan Data (CRISP-DM). CRISP-DM ialah teknik paling popular untuk projek perlombongan data, analitik dan sains data dan pertama kali diterbitkan pada tahun 1999 untuk menyeragamkan kaedah perlombongan data merentas industri (Chapman et al., 2000). Enam fasa membentuk model kitaran hayat, enam fasa tersebut ialah Pemahaman Perniagaan, Pemahaman Data, Penyediaan Data, Pemodelan, Penilaian dan Penggunaan. Ini boleh ditunjukkan dalam Rajah 1 di bawah:



Rajah 1: Fasa CRISP-DM. Sumber: (Hotz, 2022)

Fasa pemahaman terhadap perniagaan ialah langkah pertama CRISP-DM. Fasa ini membantu memastikan projek perlombongan data difokuskan untuk memenuhi keperluan dan objektif perniagaan atau organisasi. Matlamat kajian ini adalah untuk membangunkan model pembelajaran mesin menggunakan data pasaran Bitcoin terkini untuk ramalan siri masa Bitcoin. Pihak berkepentingan projek ini mungkin termasuk pelabur, institusi kewangan dan peniaga.

Fasa pemahaman data ialah langkah kedua CRISP-DM, dan ia tertumpu pada penerokaan dan menganalisis data yang dikumpul untuk mendapatkan pemahaman yang lebih baik tentang kandungan, struktur dan kualitinya. Sumber data yang digunakan untuk penyelidikan ini adalah data sejarah Bitcoin dalam mata wang Dolar Amerika Syarikat yang diperoleh daripada laman web Yahoo Finance, sebanyak 1827 rekod harga harian Bitcoin dalam tempoh dari 7 Mei 2018 hingga 7 Mei 2023 telah diekstrak untuk kajian ini. Kajian ini secara khusus memfokuskan pada analisis siri masa univariat, yang melibatkan menganalisis dan meramalkan harga Bitcoin berdasarkan nilai harga tutup.

Fasa penyediaan data ialah fasa yang melibatkan pemprosesan data untuk menghasilkan set data bersih yang sesuai untuk penyepaduan ke dalam pembelajaran mesin model. Penyediaan Data ialah langkah penting dalam proses pembelajaran mesin, kerana ia membantu memastikan data bersih dan konsisten, yang diperlukan untuk membina model yang tepat. Terlebih dahulu, penerokaan data akan dilakukan untuk memastikan pengkaji mempunyai pemahaman yang jelas tentang set data yang digunakan. Laporan kualiti data dilakukan dengan menyemak dimensi, jenis dan statistik deskriptif data. Visualisasi data juga dilakukan untuk membantu memahami corak dan arah aliran dalam set data, Matplotlib dan Plotly iaitu perpustakaan dalam python telah digunakan khas untuk tujuan visualisasi. Pembersihan data untuk kajian ini tidak perlu dilaksanakan kerana set data yang diperoleh bebas dari nilai kosong dan nilai yang sama. Namun, penskalaan data dengan Penskala Standard telah

dilakukan untuk memastikan layanan yang adil terhadap ciri dan meningkatkan ketepatan dan kebolehpercayaan keseluruhan model pembelajaran mesin. Akhir sekali, teknik tingkap gelongsor (*Sliding Window Technique*) digunakan untuk ramalan univariate iaitu dengan melibatkan membahagikan data siri masa kepada tettingkap bertindih dengan saiz tetap, di mana setiap tettingkap terdiri daripada jujukan nilai harga Tutup berturut-turut.

Dalam fasa pemodelan, model pembelajaran mesin dibangunkan dan diuji untuk mengenal pasti corak dan perhubungan dalam data. Ini biasanya dilakukan menggunakan subset data yang dikenali sebagai set latihan. Model tersebut kemudiannya diuji pada subset berasingan data yang dikenali sebagai set ujian untuk menilai keberkesannya. Model yang akan digunakan dalam penyelidikan ini ialah Mesin Vektor Sokongan Regresi (SVR), K - Jiran Terdekat Regresi (KNN), Peningkatan Kecerunan Melampau (XGBoost), dan Memori Jangka Pendek Panjang (LSTM). Penalaan hiperparameter juga dilakukan untuk mendapatkan hiperparameter yang optimum bagi setiap model pembelajaran mesin. Mesin Vektor Sokongan (SVM) ialah sejenis algoritma pembelajaran mesin yang diselia yang boleh digunakan untuk ramalan siri masa. Walau bagaimanapun, kajian ini akan menggunakan Regresi Vektor Sokongan (SVR). SVR ialah sejenis algoritma SVM yang boleh digunakan untuk tugas regresi. Seperti SVM lain, ia adalah algoritma yang mencari hyperplane dalam ruang N-dimensi yang secara maksimum memisahkan titik data satu kelas daripada kelas lain. Contoh penggunaan SVR dalam penggunaan ramalan harga Bitcoin didapati dalam kajian (Murray et al., 2023). K- Jiran Terdekat Regresi (KNN) ialah sejenis algoritma pembelajaran mesin berasaskan contoh bukan parametrik yang boleh digunakan untuk ramalan siri masa. KNN berfungsi dengan menyimpan semua data latihan dan membuat ramalan untuk titik data baharu dengan mencari K jiran terdekat bagi titik data baharu dan membuat purata nilainya (Cover & Hart, 1967). Contoh penggunaan KNN dalam penggunaan ramalan harga Bitcoin didapati dalam kajian (Murray et al., 2023). XGBoost ialah pelaksanaan peningkatan kecerunan yang direka khusus

untuk mengendalikan set data yang besar dan mengoptimumkan prestasi model. Peningkatan kecerunan ialah teknik pembelajaran mesin untuk masalah regresi, yang menghasilkan model ramalan dalam bentuk himpunan model ramalan yang lemah, biasanya pokok keputusan. Idea asas di sebalik peningkatan kecerunan adalah untuk melatih model lemah secara berurutan, masing-masing cuba membetulkan kesilapan model sebelumnya. Memori Jangka Pendek Panjang (LSTM) ialah sejenis rangkaian neural berulang (RNN) yang sangat sesuai untuk pemodelan data siri masa. RNN ialah sejenis rangkaian saraf yang mampu memproses data berjujukan, seperti siri masa, dengan menggunakan sambungan antara unit tersembunyi yang membolehkan maklumat dihantar dari langkah satu masa ke langkah seterusnya. LSTM ialah varian RNN yang mampu menangkap kebergantungan jangka panjang dalam data siri masa dengan menggunakan jenis sel memori khas dan mekanisme gating untuk mengawal aliran maklumat melalui rangkaian (Hochreiter & Schmidhuber, 1997). Contoh penggunaan LSTM dalam penggunaan ramalan harga Bitcoin didapati dalam kajian (McNally et al., 2018).

Fasa penilaian ialah di mana prestasi model terlatih diukur pada set data ujian. Set data ujian ialah set data berasingan daripada set data latihan yang tidak digunakan semasa latihan model. Tujuan fasa penilaian adalah untuk menilai keupayaan generalisasi model, iaitu keupayaannya untuk membuat ramalan yang tepat ke atas data baharu yang tidak kelihatan. Metrik penilaian yang akan digunakan untuk mengukur prestasi model ramalan siri masa bitcoin ialah Root Mean Squared Ralat (RMSE), Min Ralat Mutlak (MAE) dan R2 Skor. (Hyndman & Koehler, 2006).

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}} \quad \dots(1)$$

$$MAE = \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{n} \quad \dots(2)$$

$$R^2 = \frac{SS_{RES}}{SS_{TOT}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad \dots(3)$$

RMSE mengukur purata perbezaan kuasa dua antara nilai ramalan dan nilai sebenar dan kemudian mengambil punca kuasa dua min perbezaan kuasa dua ini, di mana persamaan 1 diatas menunjukkan "diramalkan" ( $\hat{y}_i$ ) dan "benar" ( $y_i$ ) ialah nilai yang diramalkan dan nilai sebenar, masing-masing, dan "min" dan "sqrt" ialah fungsi min dan punca kuasa dua. MAE mengukur purata perbezaan mutlak antara nilai yang diramalkan dan nilai sebenar, di mana persamaan 2 menunjukkan "diramalkan" ( $\hat{y}_i$ ) dan "benar" ( $y_i$ ) ialah nilai yang diramalkan dan nilai sebenar, masing-masing, dan "min" dan "abs" ialah fungsi nilai min dan mutlak. R kuasa dua ( $R^2$  Skor), juga dikenali sebagai pekali penentuan, ialah ukuran statistik yang digunakan untuk menilai prestasi model regresi. Ia mewakili bahagian varians dalam pembolehubah bersandar (dalam kes ini, harga Bitcoin) yang boleh diramal daripada pembolehubah bebas (peramal atau ciri) yang digunakan dalam model, di mana persamaan 3 menunjukkan SS res ialah jumlah kuasa dua baki (atau jumlah ralat kuasa dua) - perbezaan antara nilai sebenar dan nilai ramalan. SS tot ialah jumlah jumlah kuasa dua, yang mewakili jumlah kebolehubahan pembolehubah bersandar.

Fasa penggunaan ialah langkah terakhir dalam proses CRISP-DM. Matlamat fasa penggunaan adalah untuk memastikan projek ini memberikan nilai kepada organisasi dengan menyediakan model ramalan yang boleh digunakan untuk menyokong keputusan perniagaan. Kajian ini juga dijalankan untuk tujuan meningkatkan kefahaman dan dijadikan rujukan untuk kerja masa hadapan.



## Keputusan dan Perbincangan

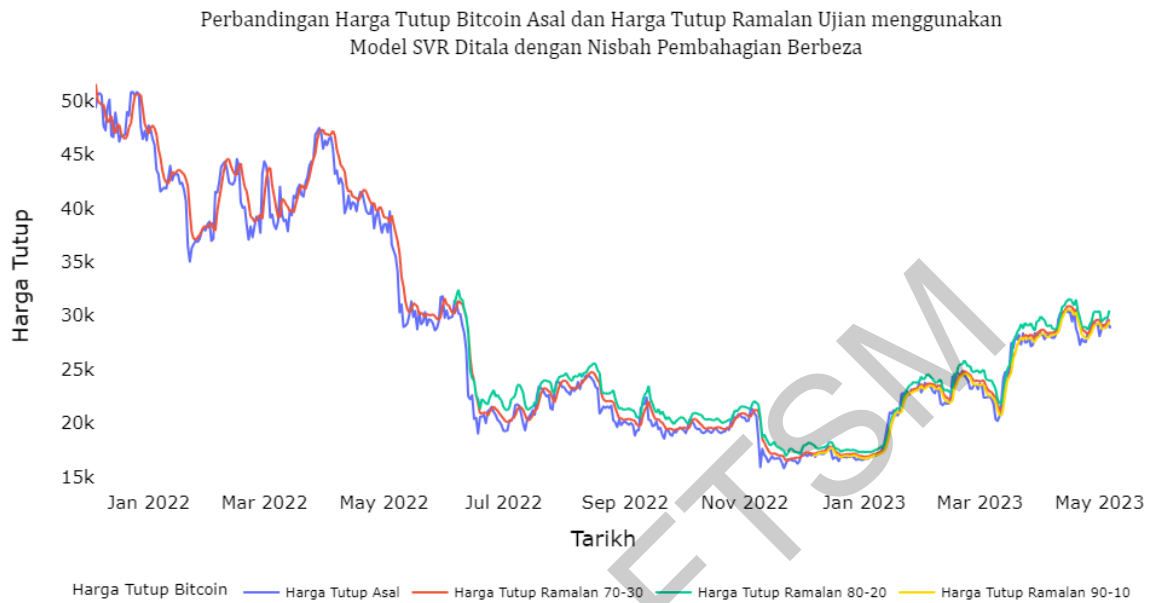
Bahagian ini membentangkan hasil penilaian metrik untuk model ramalan, membandingkan nisbah pembahagian berbeza serta dengan dan tanpa penalaan hiperparameter, dipaparkan dalam bentuk jadual dan graf siri masa.

### Penilaian Pemodelan Regresi Vektor Sokongan (SVR)

Jadual 1: Keputusan penilaian metrik model ramalan SVR yang telah ditala

Nisbah Pembahagian (SVR)	Penalaan	Penilaian Metriks					
		RMSE (Latih)	RMSE (Ujian)	MAE (Latih)	MAE (Ujian)	R2 Skor (Latih)	R2 Skor (Ujian)
70 - 30	Tidak Ditala	1185.63	1914.77	830.08	1602.95	0.99	0.96
	Ditala	1256.18	1291.70	884.26	902.82	0.99	0.98
80 - 20	Tidak Ditala	1266.41	2132.20	886.73	1847.03	0.99	0.70
	Ditala	1120.96	1381.89	755.01	1167.09	0.99	0.87
90 - 10	Tidak Ditala	1226.70	879.69	854.73	700.89	0.99	0.96
	Ditala	1304.65	727.37	960.38	493.35	0.99	0.97

Jadual 1 menunjukkan prestasi model SVR untuk ramalan harga bitcoin dengan dan tanpa penalaan. Penalaan model secara amnya membawa kepada peningkatan dalam ketepatan ramalan. Model yang ditala mencapai nilai RMSE dan MAE yang lebih rendah, menunjukkan pengurangan ralat dalam meramalkan harga bitcoin. Secara ketara, nisbah pembahagian 80-20 menunjukkan peningkatan paling ketara selepas penalaan, dengan RMSE menurun daripada 1266.41 kepada 1120.96, MAE menurun daripada 886.73 kepada 755.01 untuk data ujian dan skor R2 meningkat daripada 0.70 kepada 0.87. Nisbah pembahagian 90-10 menunjukkan prestasi yang lebih baik selepas penalaan, mencapai penurunan RMSE terendah daripada 879.69 kepada 727.37 dan MAE menurun daripada 700.89 kepada 493.35 untuk data ujian. Nisbah pembahagian 70-30 juga terdapat peningkatan dalam RMSE dan MAE juga dengan skor R2 mencapai tertinggi pada 0.98. Oleh itu, penalaan model SVR boleh meningkatkan prestasinya dengan ketara dan menjadikannya lebih tepat dalam meramalkan harga Bitcoin. Dalam kes ini, nisbah pembahagian 90-10 nampaknya memberikan prestasi model SVR terbaik pada set data ujian berdasarkan metrik penilaian yang diberikan, kerana ia mempamerkan ralat ramalan terendah (RMSE dan MAE) dan skor R2 yang agak tinggi.



Graf 1: Graf keputusan harga tutup Bitcoin ramalan ujian menggunakan model SVR ditala dengan nisbah pembahagian berbeza

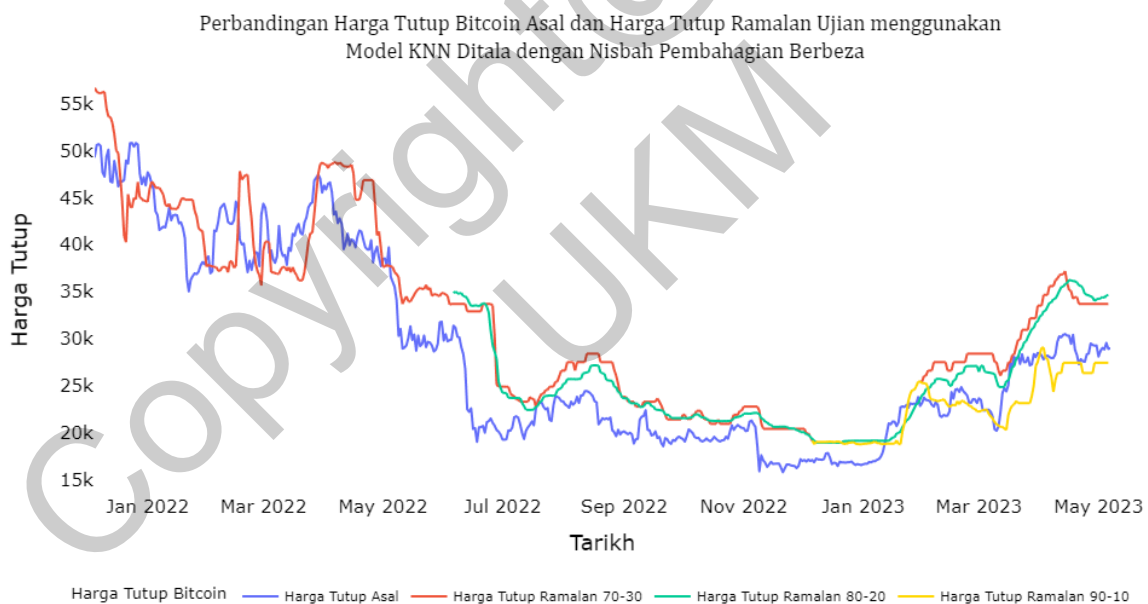
Berdasarkan Graf 1 iaitu graf siri masa, dapat dilihat bahawa pembahagian 80-20 menunjukkan peningkatan dalam penjajaran antara harga yang diramalkan dan harga sebenar berbanding graf siri masa model SVR yang tidak ditala. Bagi pembahagian 70-30 dan 90-10, ia masih sesuai dan menunjukkan persamaan yang rapat antara harga yang diramalkan dan sebenar, dengan sisihan minimum.

### Penilaian Pemodelan K-Jiran Terdekat Regresi (KNN)

Jadual 2: Keputusan penilaian metrik model ramalan KNN yang telah ditala

Nisbah Pembahagian (KNN)	Penalaan	Penilaian Metriks					
		RMSE (Latih)	RMSE (Ujian)	MAE (Latih)	MAE (Ujian)	R2 Skor (Latih)	R2 Skor (Ujian)
70 - 30	Tidak Ditala	688.38	4937.51	369.65	4223.59	0.99	0.76
	Ditala	1093.16	4631.46	577.90	3938.92	0.99	0.79
80 - 20	Tidak Ditala	754.85	4758.47	434.53	4061.35	0.99	-0.50
	Ditala	0.0038	4175.52	0.0010	3347.56	0.99	-0.16
90 - 10	Tidak Ditala	743.42	2272.56	432.66	1912.36	0.99	0.74
	Ditala	1176.24	2228.91	679.87	1886.91	0.99	0.75

Berdasarkan Jadual 2, model KNN yang ditala mencapai nilai RMSE dan MAE yang lebih rendah, menunjukkan pengurangan ralat dalam meramalkan harga bitcoin. Nisbah pembahagian 90-10 menunjukkan prestasi yang lebih baik selepas penalaan, mencapai RMSE terendah (2228.91) dan MAE (1886.91) untuk data ujian. Walaupun RMSE pada data latihan nisbah 90-10 dan 70-30 meningkat dan mungkin kelihatan berlawanan dengan intuitif, ia sebenarnya boleh menjadi tanda generalisasi yang lebih baik. Ia menunjukkan bahawa model itu menjadi kurang overfit dengan set latihan dan menangkap corak yang lebih bermakna yang boleh digunakan pada data yang baharu. Oleh itu, penalaan parameter terbukti dapat meningkatkan prestasi model KNN. Dalam kes ini, nisbah pembahagian 90-10 nampaknya memberikan prestasi model terbaik pada dataset ujian berdasarkan metrik penilaian yang diberikan, kerana ia mempamerkan RMSE dan MAE terendah dan skor R2 yang agak baik (0.75).



Graf 2: Graf keputusan harga tutup Bitcoin ramalan ujian menggunakan model KNN ditala dengan nisbah pembahagian berbeza

Berdasarkan Graf 2, dapat dilihat garis ramalan berhenti bertindih antara satu sama lain selepas model ditala. Pembahagian 80-20 menunjukkan peningkatan di mana garis ramalan lebih hampir ke arah garis harga sebenar, tidak seperti pembahagian 70-30 di mana garis ramalan tidak banyak

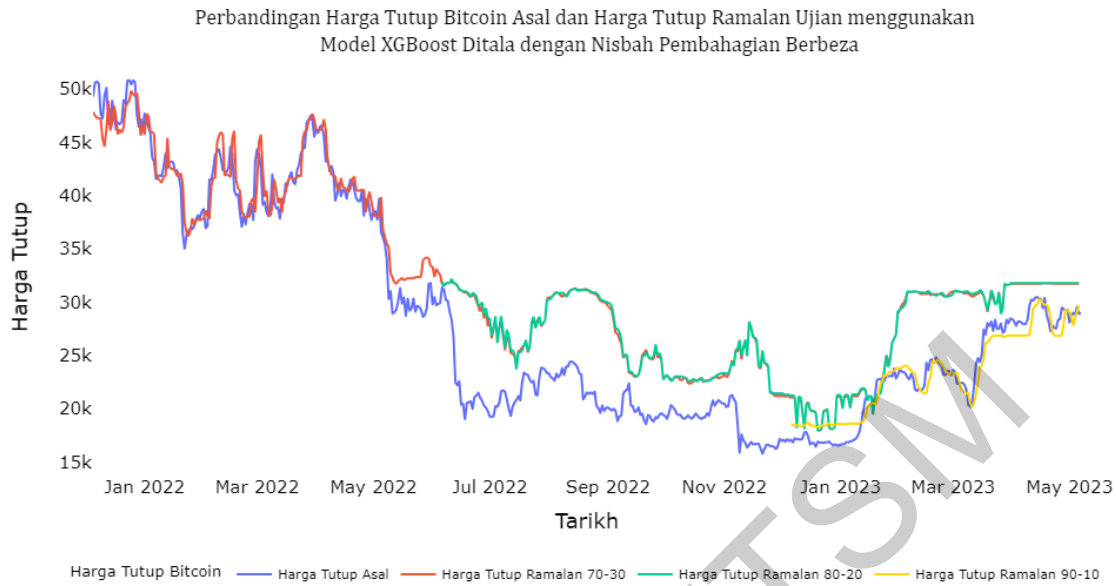
perbezaan berbanding graf siri masa KNN yang tidak ditala. Walau bagaimanapun, pembahagian 90-10 juga menunjukkan penambahbaikan di mana terdapat sedikit ralat berbanding model yang tidak ditala, terutamanya pada awal Mei 2023.

### Penilaian Pemodelan Peningkatan Kecerunan Melampau (XGBoost)

Jadual 3: Keputusan penilaian metrik model ramalan XGBoost yang telah ditala

Nisbah Pembahagian (XGBoost)	Penalaan	Penilaian Metriks					
		RMSE (Latih)	RMSE (Ujian)	MAE (Latih)	MAE (Ujian)	R2 Skor (Latih)	R2 Skor (Ujian)
70 - 30	Tidak Ditala	41.43	5851.85	28.87	4646.62	0.99	0.66
	Ditala	633.912	4860.66	343.38	3860.80	0.99	0.77
80 - 20	Tidak Ditala	58.42	7438.66	39.97	6497.72	0.99	-2.67
	Ditala	734.95	5943.47	412.72	5214.26	0.99	-1.34
90 - 10	Tidak Ditala	80.39	1547.61	54.38	1267.09	0.99	0.88
	Ditala	770.32	1222.05	457.17	1026.77	0.99	0.93

Jadual 3 membentangkan prestasi model XGBoost yang ditala untuk meramalkan harga Bitcoin. Sebelum penalaan, model menunjukkan keputusan sederhana, dengan nilai RMSE antara 41.43 hingga 80.39 pada data latihan, dan nilai yang lebih tinggi antara 1222.05 dan 7438.66 pada data ujian. Selepas penalaan, prestasi model meningkat dengan ketara, mencapai nilai RMSE yang lebih rendah pada kedua-dua data latihan dan ujian. Sebagai contoh, dalam pembahagian 70-30, model yang ditala mencapai RMSE sebanyak 633.912 pada data latihan dan 4860.66 pada data ujian, menunjukkan pengurangan yang ketara dalam ralat ramalan. Nilai MAE mengikut corak yang sama, dengan penurunan ketara selepas penalaan. Selain itu, skor R2 untuk model yang ditala kekal tinggi secara konsisten, terutamanya pembahagian 90-10 (0.93), menunjukkan kesesuaian yang baik dengan data dan ketepatan ramalan yang lebih baik. Keputusan ini menunjukkan bahawa proses penalaan meningkatkan prestasi model, membawa kepada ramalan harga Bitcoin yang lebih tepat.



Graf 3: Graf keputusan harga tutup Bitcoin ramalan ujian menggunakan model XGBoost ditala dengan nisbah pembahagian berbeza

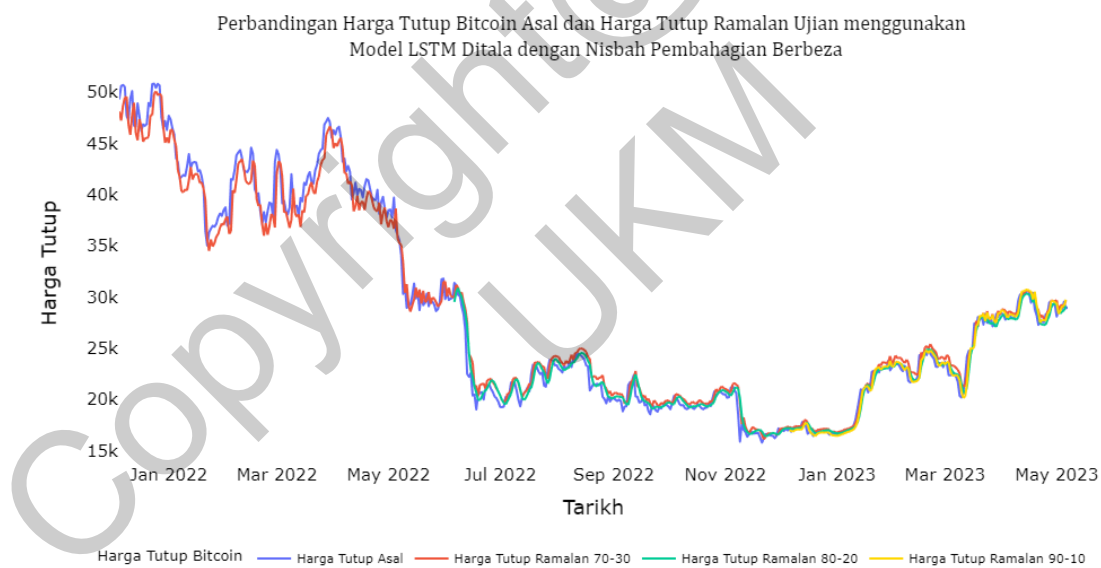
Berdasarkan Graf 3 iaitu graf siri masa, dapat dilihat bahawa jurang antara harga sebenar dan garis ramalan untuk pembahagian 70-30 juga pembahagian 80-20 hanya mempunyai peningkatan yang minimum selepas penalaan. Sebaliknya, ramalan pecahan 90-10 bertambah baik dan menunjukkan kesesuaian terbaik berbanding yang lain di mana ia sejajar dengan lebih dekat ke arah garis harga sebenar.

### Penilaian Pemodelan Memori Jangka Pendek Panjang (LSTM)

Jadual 4: Keputusan penilaian metrik model ramalan LSTM yang telah ditala

Nisbah Pembahagian (LSTM)	Penalaan	Penilaian Metriks					
		RMSE (Latih)	RMSE (Ujian)	MAE (Latih)	MAE (Ujian)	R2 Skor (Latih)	R2 Skor (Ujian)
70 - 30	Tidak Ditala	3391.44	2122.01	2414.32	1525.41	0.96	0.95
	Ditala	1130.55	1215.43	590.09	889.51	0.99	0.98
80 - 20	Tidak Ditala	2732.21	1950.05	1828.37	1373.63	0.97	0.74
	Ditala	1314.18	821.58	799.92	555.46	0.99	0.95
90 - 10	Tidak Ditala	2856.42	1908.52	1910.56	1485.41	0.97	0.81
	Ditala	1162.97	661.66	745.57	462.63	0.99	0.98

Berdasarkan Jadual 4 yang menunjukkan keputusan ditala model LSTM, model yang tidak ditala pada mulanya menghasilkan keputusan yang baik, dengan nilai RMSE antara 2732.21 hingga 3391.44 pada data latihan dan 1908.52 hingga 2122.01 pada data ujian. Selepas penalaan, prestasi model bertambah baik dengan ketara. Sebagai contoh, dalam pembahagian 70-30, model LSTM yang ditala mencapai RMSE sebanyak 1130.55 pada data latihan dan 1215.43 pada data ujian, menunjukkan pengurangan ketara dalam ralat ramalan. Penambahbaikan yang sama diperhatikan merentas nisbah pisah yang lain, terutamanya nisbah pisah 90-10 yang mempunyai RMSE terendah (661.66) dan MAE (462.63) pada data ujian. Nilai MAE dan skor R2 juga menunjukkan peningkatan ketara selepas penalaan, seterusnya mengesahkan ketepatan model yang lebih baik dan sesuai dengan data. Dalam kes ini, pembahagian 90-10 nampaknya paling menjanjikan kerana ia mempunyai ralat ramalan yang rendah dan skor R2 yang tinggi (0.98).



Graf 4: Graf keputusan harga tutup Bitcoin ramalan ujian menggunakan model LSTM ditala dengan nisbah pembahagian berbeza

Berdasarkan Graf 4, dapat dilihat dengan jelas bahawa semua ramalan nisbah pembahagian menunjukkan prestasi yang sangat baik dengan mempunyai kesesuaian yang ketat dan persamaan yang hampir dengan harga sebenar, mengesan turun naiknya dengan ketepatan yang luar biasa.

Dengan membandingkannya dengan graf siri masa model LSTM yang belum ditala, sisihan telah berkurang dalam margin yang besar. Ketepatan yang baik ini mencerminkan keberkesanan seni bina LSTM dalam memahami corak dan kebergantungan asas dalam data siri masa.

### **Perbandingan Model Secara Keseluruhan**

Penyelidik bertujuan untuk membandingkan prestasi empat model berbeza untuk ramalan harga mata wang Bitcoin: SVR, KNN, XGBoost dan LSTM. Untuk menilai model dengan tepat, penyelidik mempertimbangkan tiga nisbah pembahagian data latih-uji yang berbeza: 70-30, 80-20, dan 90-10, lalu menjalankan penyiasatan terhadap kesan penalaan hiperparameter pada prestasi model, membandingkan keputusan mereka sebelum dan selepas penalaan.

Selepas menganalisis keputusan yang diperolehi, ternyata bahawa nisbah pembahagian data latih-uji 90-10 secara konsisten mengatasi nisbah lain merentas semua model. Dapatan ini menunjukkan bahawa memperuntukkan sebahagian besar data untuk latihan (90%) dan bahagian yang lebih kecil untuk ujian (10%) menghasilkan ketepatan ramalan yang lebih baik. Selain itu, kesemua model tersebut menunjukkan peningkatan yang ketara dalam prestasi selepas menjalani penalaan hiperparameter. Ini menandakan kepentingan memilih hiperparameter optimum dengan teliti untuk meningkatkan keupayaan ramalan model. Memandangkan penemuan ini, penyelidik membuat keputusan termaklum untuk meneruskan eksperimen dengan nisbah menggunakan nisbah pembahagian 90-10 dan model yang ditala untuk eksperimen penyelidik yang seterusnya. Jadual 4.15 di bawah menunjukkan perbandingan antara nisbah pembahagian 90-10 model ditala.

Jadual 5: Perbandingan model ramalan yangtelah ditala (nisbah pembahagian 90-10)

Model	Penalaan	Penilaian Metriks						Ranking
		RMSE (Latih)	RMSE (Ujian)	MAE (Latih)	MAE (Ujian)	R2 Skor (Latih)	R2 Skor (Ujian)	
SVR	{C: 10, Gamma: 0.001, Epsilon: 0.1}	1304.65	727.37	960.38	493.35	0.99	0.97	2
KNN	{n-neighbors: 10, Weight: Uniform}	1176.24	2228.91	679.87	1886.91	0.99	0.75	4
XGBOOST	{Colsample Bytree: 0.5, Learning rate: 0.01, MaxDepth: 3, n-estimators: 1000}	770.32	1222.05	457.17	1026.77	0.99	0.93	3
LSTM	{Learning rate: 0.001. Units: 64}	1162.97	661.66	745.57	462.63	0.99	0.98	1

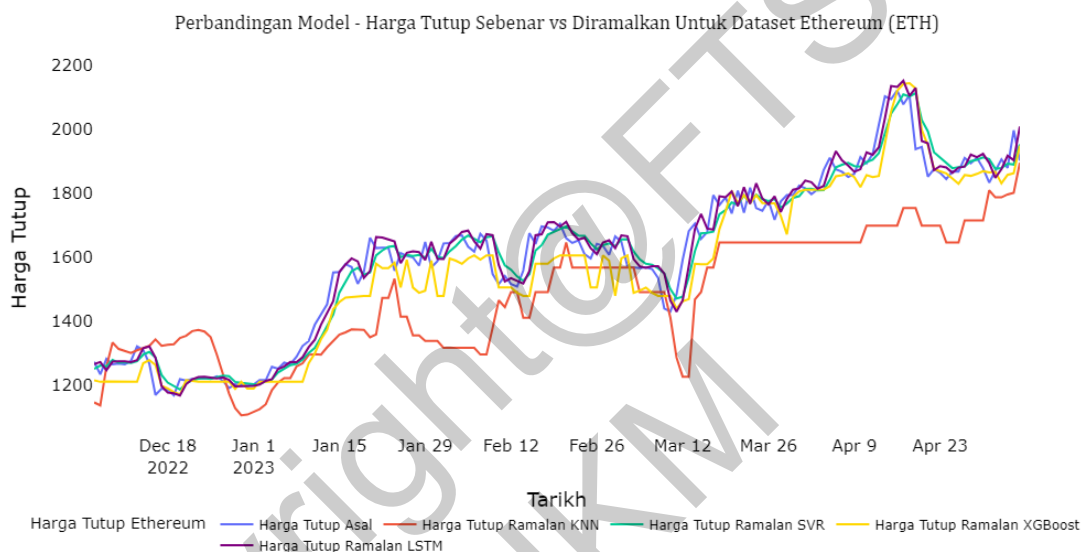
Berdasarkan Jadual 5 di atas, antara empat model yang dinilai untuk ramalan siri masa mata wang Bitcoin, SVR, XGBoost dan LSTM menunjukkan prestasi yang kukuh, manakala KNN mempamerkan keputusan yang agak lemah. SVR mempamerkan ramalan tepat dengan nilai RMSE dan MAE yang rendah, menangkap kedua-dua turun naik jangka pendek dan arah aliran jangka panjang. XGBoost menangkap corak dengan berkesan dengan nilai RMSE dan MAE yang rendah, menunjukkan ramalan yang tepat dan hubungan linear yang kukuh dengan harga sebenar. LSTM, sebagai rangkaian saraf berulang, cemerlang dalam memodelkan kebergantungan temporal, menghasilkan nilai RMSE dan MAE yang rendah, dan skor R2 yang tinggi. Sebaliknya, KNN memaparkan ralat yang lebih tinggi dan skor R2 yang lebih rendah, menggambarkan batasan dalam menangkap kerumitan dinamik harga mata wang Bitcoin. Secara keseluruhan, model-model tersebut telah disenaraikan berdasarkan Skor R2, yang menghasilkan LSTM sebagai tempat pertama, SVR kedua, XGBoost ketiga dan terakhir KNN. Model ini telah disimpan dan akan disuaikan untuk menilai prestasi model dengan menggunakan data mata wang kripto yang lain.



## Kebolegunaan model pada set data Ethereum (ETH)

Jadual 6: Keputusan penilaian metrik model ramalan untuk set data Ethereum (ETH)

Model	RMSE (Train)	RMSE (Test)	MAE (Train)	MAE (Test)	R2 Skor (Train)	R2 Skor (Test)	Ranking
KNN	169.82	183.16	97.96	153.10	0.98	0.46	4
SVR	101.53	51.27	72.88	37.52	0.99	0.95	2
XGBOOST	103.76	74.92	55.84	58.46	0.99	0.91	3
LSTM	89.66	48.29	54.46	35.04	0.99	0.96	1



Graf 5: Graf harga tutup sebenar vs ramalan untuk dataset Ethereum (ETH)

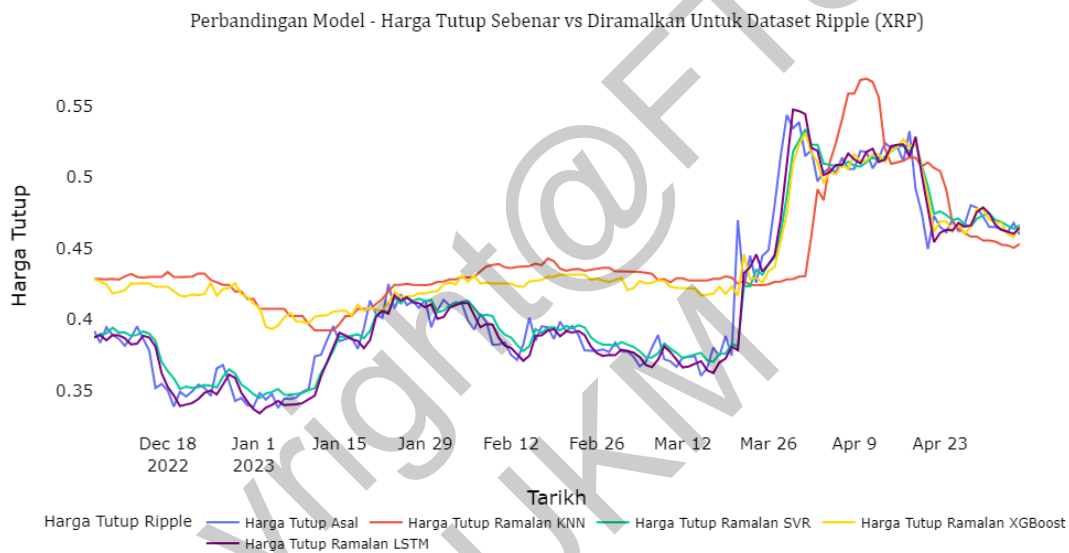
Berdasarkan Jadual 6, antara model yang dinilai untuk ramalan harga Ethereum (ETH), LSTM muncul sebagai prestasi terbaik, mencapai nilai RMSE (48.29) dan MAE (54.46) terendah, menunjukkan ramalan yang tepat dengan ralat yang minimum. LSTM menunjukkan keupayaan ramalan yang sangat baik dengan skor R2 yang tinggi sebanyak 0.96, dengan berkesan menangkap corak dan arah aliran dalam data harga. SVR mempamerkan prestasi kompetitif dengan nilai RMSE rendah (51.27) dan MAE (72.88), bersama-sama dengan skor R2 yang tinggi sebanyak 0.95. XGBoost menunjukkan prestasi yang baik, walaupun dengan metrik ralat yang lebih tinggi sedikit, manakala KNN menunjukkan ralat yang lebih tinggi dan skor R2 yang agak rendah iaitu 0.46. Secara

keseluruhan, LSTM dan SVR terbukti sebagai model yang paling sesuai untuk ramalan harga ETH, memberikan kesesuaian yang ketat dan hampir menyerupai harga ETH sebenar dalam Graf 5.

### Kebolegunaan model pada set data Ripple (XRP)

Jadual 7: Keputusan penilaian metrik model ramalan untuk set data Ripple (XRP)

Model	RMSE (Train)	RMSE (Test)	MAE (Train)	MAE (Test)	R2 Skor (Train)	R2 Skor (Test)	Ranking
KNN	0.09388	0.04915	0.05215	0.04203	0.90	0.23	4
SVR	0.04550	0.01487	0.02495	0.00967	0.98	0.93	2
XGBOOST	0.06678	0.03958	0.03267	0.03265	0.95	0.50	3
LSTM	0.04150	0.01404	0.02179	0.00933	0.98	0.94	1



Graf 6: Graf harga tutup sebenar vs ramalan untuk dataset Ripple (XRP)

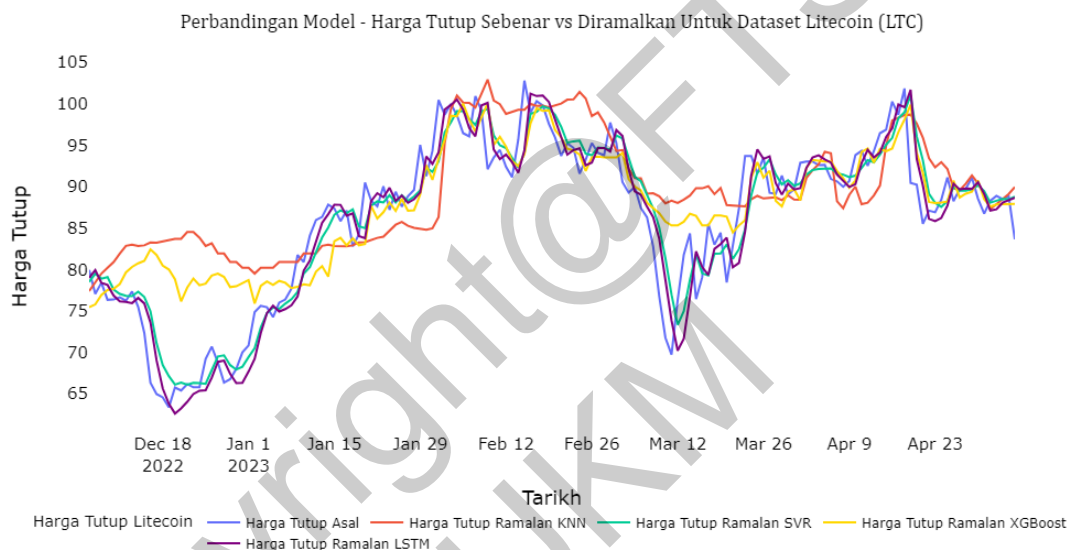
Berdasarkan Jadual 7, untuk ramalan harga Ripple (XRP), LSTM muncul sebagai model berprestasi tinggi, dengan RMSE terendah (0.01404) dan MAE (0.00933), menunjukkan ramalan yang sangat tepat. LSTM mempamerkan koresponden yang kukuh dengan harga sebenar, dicerminkan dalam skor R2 yang tinggi sebanyak 0.94. SVR memaparkan prestasi kompetitif dengan RMSE rendah (0.01487) dan MAE (0.00967), bersama-sama dengan skor R2 yang tinggi sebanyak 0.93. XGBoost dan KNN menunjukkan ralat yang lebih tinggi dan skor R2 yang agak rendah (0.50 dan 0.23). Secara keseluruhan, LSTM dan SVR terbukti sebagai model yang paling sesuai untuk ramalan harga XRP.

Graf 6 menggambarkan prestasi kukuh LSTM dan SVR yang hampir menyerupai garis harga sebenar XRP, manakala XGBoost dan KNN menunjukkan sisihan yang lebih besar.

### Kebolegunaan model pada set data Litecoin (LTC)

Jadual 8: Keputusan penilaian metrik model ramalan untuk set data Litecoin (LTC)

Model	RMSE (Train)	RMSE (Test)	MAE (Train)	MAE (Test)	R2 Skor (Train)	R2 Skor (Test)	Ranking
KNN	16.07077	7.48355	9.23912	5.64029	0.92	0.41	4
SVR	8.08788	3.33284	4.76580	2.46509	0.98	0.88	2
XGBOOST	13.22260	5.61960	6.08134	3.95008	0.95	0.69	3
LSTM	7.40682	3.26842	4.16039	2.41940	0.98	0.89	1



Graf 7: Graf harga tutup sebenar vs ramalan untuk dataset Litecoin (LTC)

Jadual 8 menunjukkan penilaian metrik ramalan harga Litecoin (LTC), dan sekali lagi LSTM muncul sebagai prestasi terbaik dengan RMSE terendah (3.26842) dan MAE (2.41940), menunjukkan ramalan yang tepat. LSTM mencapai markah R2 tertinggi (0.89). SVR memaparkan prestasi kompetitif dengan RMSE rendah (3.33284) dan MAE (2.46509), bersama-sama dengan skor R2 yang tinggi (0.88). XGBoost berprestasi secara sederhana dengan metrik ralat yang lebih tinggi sedikit dan skor R2 yang lebih rendah sebanyak 0.69, perkara yang sama berlaku kepada KNN di mana skor R2 ialah 0.41. Secara keseluruhan, LSTM dan SVR juga merupakan model yang paling sesuai untuk ramalan harga LTC, dengan LSTM menunjukkan prestasi terbaik. Berdasarkan Graf 7, sama seperti

set data XRP, LSTM dan SVR mempamerkan ketepatan yang baik, kerana garis harga dekat LTC yang diramalkan menjejaki garis harga sebenar dengan rapat.

### **Perbincangan Kebolehgunaan Model Secara Keseluruhan**

Selepas menilai prestasi model pada tiga mata wang kripto yang berbeza (Ethereum, Ripple, dan Litecoin) untuk ramalan harga, beberapa pemerhatian boleh dibuat. LSTM secara konsisten muncul sebagai model berprestasi tinggi merentas ketiga-tiga mata wang kripto, memaparkan ketepatan yang luar biasa dalam menangkap dinamik harga. Ia mencapai nilai RMSE dan MAE terendah, menunjukkan ramalan yang tepat dengan ralat yang minimum. Selain itu, LSTM menunjukkan skor R2 yang tinggi, mencerminkan korespondensi yang kukuh antara harga yang diramalkan dan harga sebenar.

SVR juga mempamerkan prestasi kompetitif, dengan metrik ralat yang agak rendah dan skor R2 yang tinggi, menjadikannya pilihan yang sesuai untuk ramalan harga. XGBoost memaparkan prestasi yang baik, walaupun dengan metrik ralat yang lebih tinggi sedikit berbanding LSTM dan SVR. KNN, sebaliknya, mempamerkan ralat yang lebih tinggi dan skor R2 yang lebih rendah, mencadangkan batasan dalam menangkap kerumitan corak harga mata wang kripto.

Keteguhan model yang digunakan untuk ramalan harga mata wang kripto terbukti daripada prestasi konsisten mereka merentas berbilang set data. LSTM secara konsisten menduduki tempat pertama, SVR kedua, XGBoost ketiga dan KNN terakhir untuk keempat-empat set data, termasuk Bitcoin, Ethereum, Ripple dan Litecoin. Ketekalan dalam kedudukan ini menunjukkan keupayaan model untuk membuat generalisasi dengan baik dan menangkap corak dan aliran asas dalam data harga mata wang kripto yang berbeza.

Hakikat bahawa LSTM dan SVR secara konsisten disenaraikan dalam kalangan yang berprestasi terbaik merentas semua set data, menunjukkan keteguhan dan kesesuaian dalam tugas ramalan harga mata wang kripto. Ini mencadangkan bahawa model tersebut menghasilkan prestasi

yang konsisten merentas pelbagai mata wang kripto, mengukuhkan kebolehpercayaan dan keberkesannya dalam menangkap dinamik kompleks pasaran mata wang kripto.

Kesimpulannya, penyelidikan ke atas ramalan siri masa harga Bitcoin menggunakan pelbagai model pembelajaran mesin dan nisbah pembahagian data latih-uji yang berbeza mendedahkan pandangan dan pemahaman yang mendalam. Pilihan nisbah pembahagian sememangnya memberi kesan kepada prestasi ramalan, dengan pembahagian 90-10 muncul sebagai yang paling sesuai untuk semua model. Selain itu, keputusan ramalan jelas menunjukkan kepentingan penalaan hiperparameter, kerana model yang ditala secara konsisten mengatasi prestasi yang tidak ditala. Antara model yang dibandingkan, LSTM mempamerkan ketepatan ramalan tertinggi, diikuti oleh SVR dan XGBoost, manakala KNN Regressor ketinggalan sedikit. Penemuan ini menekankan kepentingan pembahagian data yang betul dan pengoptimuman hiperparameter dalam tugas ramalan siri masa dan menyerlahkan LSTM sebagai pilihan yang teguh untuk ramalan harga Bitcoin.

Beberapa cadangan dan kerja masa depan yang berpotensi muncul daripada projek ini. Satu cadangan yang menjanjikan ialah sistem ramalan masa nyata boleh dibangunkan untuk menawarkan pandangan terkini untuk pedagang dan pelabur, dengan mengambil kira sifat dinamik pasaran mata wang kripto. Tambahan pula, menggabungkan faktor luaran seperti perubahan peraturan, peristiwa makroekonomi atau berita mata wang kripto boleh memberi manfaat dalam menangkap sentimen pasaran dan meningkatkan ketepatan model dalam menangkap dinamik pasaran dunia sebenar.

## **Papan Pemuka**

Dalam penyelidikan ini, penyelidik menggunakan *Streamlit* untuk membangunkan papan pemuka aplikasi web yang mempamerkan hasil dan visualisasi projek. Rajah 2 mempamerkan papan pemuka yang dihasilkan menggunakan *Streamlit*. Papan pemuka tersebut mengandungi analisis set data mata wang kripto yang digunakan dalam projek ini. Melalui papan permukaan tersebut, pengguna dapat melihat nilai ramalan bersama nilai sebenar, membolehkan pengguna membandingkan secara visual

ketepatan ramalan. Selain itu, carta bar yang mewakili skor R2 untuk setiap mata wang kripto akan ditunjukkan.



Rajah 2: Papan Pemuka

## Kesimpulan

Sepanjang projek ini, penyelidik menumpukan pada ramalan siri masa harga Bitcoin menggunakan model pembelajaran mesin, iaitu Regresi Vektor Sokongan (SVR), Memori Jangka Pendek Panjang (LSTM), Peningkatan Kecerunan Melampau (XGBoost) dan model K-Jiran Terdekat Regresi (KNN). Penyelidik meneroka kesan nisbah pembahagian datalatih-ujij yang berbeza pada prestasi model dan menjalankan penalaan hiperparameter untuk meningkatkan ketepatan ramalan. Penemuan mendedahkan bahawa pilihan nisbah pembahagian secara signifikan mempengaruhi prestasi model,

dengan pembahagian 90-10 adalah yang paling sesuai untuk semua model. Selain itu, penalaan hiperparameter terbukti penting, kerana model yang ditala secara konsisten mengatasi prestasi yang tidak ditala. Antara model yang dibandingkan, LSTM menunjukkan ketepatan ramalan yang paling tinggi, diikuti oleh SVR, manakala XGBoost dan KNN menunjukkan hasil yang agak kurang berprestasi. Penyelidikan itu juga memanjangkan penyiasatannya kepada set data mata wang kripto alternatif (ETH, XRP, dan LTC), dan keputusan menunjukkan bahawa LSTM dan SVR kekal teguh dan berkesan dalam meramalkan arah aliran harga merentas aset ini.

Kesimpulannya, projek ini telah berjaya mencapai objektifnya untuk meramalkan pergerakan harga siri masa Bitcoin menggunakan pelbagai model pembelajaran mesin dan meneroka kesan nisbah pembahagian data latihan-uji yang berbeza. Penilaian komprehensif model SVR, LSTM, XGBoost dan KNN memberikan cerapan berharga tentang prestasinya, dengan LSTM muncul sebagai model paling tepat untuk ramalan harga Bitcoin. Selain itu, penyiasatan set data mata wang kripto alternatif mempamerkan lagi kebolehsuaian LSTM dan SVR. Walau bagaimanapun, projek itu juga menyerlahkan cabaran yang ditimbulkan oleh batasan pengiraan dalam platform berasaskan awan dan keperluan untuk penalaan hiperparameter yang meluas. Bagi penyelidik masa depan, projek ini berfungsi sebagai asas akademik yang kukuh untuk penyelidikan ramalan harga mata wang kripto, menawarkan panduan tentang pemilihan model, dan pengoptimuman hiperparameter untuk meningkatkan keupayaan ramalan.

## Penghargaan

Segala puji dan syukur bagi Allah, Yang Maha Pengasih lagi Maha Penyayang. Saya dengan rendah hati menadah tangan tanda kesyukuran, terharu kerana telah menyiapkan projek tahun akhir saya dengan bimbingan dan rahmat yang dikurniakan kepada saya. Saya ingin merakamkan setinggi-tinggi penghargaan kepada penyelia saya, Prof. Dr. Salwani Abdullah, atas bimbingan, sokongan dan dorongan sepanjang setahun sebagai penyelia projek akhir tahun saya. Kepakaran dan pandangan beliau yang tidak ternilai telah memainkan peranan penting dalam membentuk hala tuju dan hasil projek saya. Saya amat berterima kasih atas masa dan usaha yang dia telah laburkan untuk saya dan kerja saya. Saya tidak akan dapat menyiapkan projek ini tanpa bimbingan beliau. Terima kasih, Prof. Dr. Salwani.

Saya juga ingin mengucapkan terima kasih juga kepada para pensyarah dari FTSM yang telah menaburkan ilmu pengetahuan kepada saya sepanjang pengajian saya di Universiti Kebangsaan Malaysia (UKM). Saya juga ingin merakamkan ribuan terima kasih kepada ahli keluarga saya kerana memberi sokongan dan semangat serta sentiasa berdoa terhadap kejayaan saya di universiti.

Akhirnya, terima kasih kepada semua yang terlibat secara langsung atau tidak langsung dalam menghasilkan projek tahun akhir ini. Saya juga ingin memohon maaf sekiranya terdapat kesilapan sepanjang pelaksanaan projek akhir tahun ini. Sekian, terima kasih.



## RUJUKAN

- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T. P., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide.
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21-27. <https://doi.org/10.1109/TIT.1967.1053964>
- Fleischer, J., von Laszewski, G., Suarez, C., & Parra Bautista, Y. (2022). Time Series Analysis of Cryptocurrency Prices Using Long Short-Term Memory. *Algorithms*, 15, 230. <https://doi.org/10.3390/a15070230>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-term Memory. *Neural computation*, 9, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hotz, N. (2022). *What is CRISP DM?* <https://www.datascience-pm.com/crisp-dm-2/>
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679-688. <https://doi.org/https://doi.org/10.1016/j.ijforecast.2006.03.001>
- McNally, S., Roche, J., & Caton, S. (2018). *Predicting the Price of Bitcoin Using Machine Learning*. <https://doi.org/10.1109/PDP2018.2018.00060>
- Mudassir, M., Bennbaia, S., Unal, D., & Hammoudeh, M. (2020). Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. *Neural Computing and Applications*. <https://doi.org/10.1007/s00521-020-05129-6>
- Murray, K., Rossi, A., Carraro, D., & Visentin, A. (2023). On Forecasting Cryptocurrency Prices: A Comparison of Machine Learning, Deep Learning, and Ensembles. *Forecasting*, 5(1), 196–209. <https://doi.org/10.3390/forecast5010010>