

SKRIPSI

ANALISIS MODEL *CNN* DAN *LSTM* UNTUK MEMPREDIKSI HARGA *BITCOIN*



Gregorius Bryan Lee Haryanto

NPM: 6161801072

PROGRAM STUDI MATEMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS
UNIVERSITAS KATOLIK PARAHYANGAN
2022

FINAL PROJECT

ANALYSIS OF CNN AND LSTM MODEL FOR FORECASTING BITCOIN.



Gregorius Bryan Lee Haryanto

NPM: 6161801072

**DEPARTMENT OF MATHEMATICS
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES
PARAHYANGAN CATHOLIC UNIVERSITY
2022**

LEMBAR PENGESAHAN

ANALISIS MODEL *CNN* DAN *LSTM* UNTUK MEMPREDIKSI HARGA *BITCOIN*

Gregorius Bryan Lee Haryanto

NPM: 6161801072

Bandung, 11 Januari 2022

Menyetujui,

Pembimbing 1



Agus Sukmana, M.Sc.

Pembimbing 2



Dr. Andreas Parama Wijaya

Ketua Tim Penguji



Taufik Limansyah, M.T.

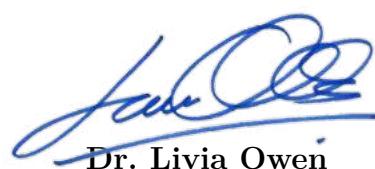
Anggota Tim Penguji



Maria Anestasia, M.Si., M.Act.Sc.

Mengetahui,

Ketua Program Studi



Dr. Livia Owen

PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

ANALISIS MODEL *CNN* DAN *LSTM* UNTUK MEMPREDIKSI HARGA BITCOIN

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuahkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,
Tanggal 11 Januari 2022



Gregorius Bryan Lee Haryanto
NPM: 6161801072

ABSTRAK

Salah satu alasan seseorang berinvestasi adalah untuk mendapatkan keuntungan. Saat ini, mata uang kripto (*cryptocurrency*) merupakan salah satu jenis investasi yang dapat digunakan. Prediksi harga mata uang kripto di masa yang akan datang dapat membantu proses pengambilan keputusan dalam berinvestasi. Salah satu cara memprediksi harga mata uang kripto di masa yang akan datang adalah dengan menggunakan sistem saraf buatan (*artificial neural network*). Terdapat dua jenis sistem saraf buatan yang digunakan, yaitu LSTM (*long short-term memory*) dan gabungan CNN (*convolutional neural network*) dan LSTM. Nilai dari error menjadi suatu tolak ukur sistem saraf buatan yang baik. Data yang digunakan terdiri dari data pada tahun 2020-2021 dengan jarak antar data adalah 1 menit, dan dibagi menjadi tiga periode, yaitu data periode Januari 2020 - Juli 2020, data periode Januari 2020 - Desember 2020, dan data periode Januari 2021 - November 2021. Data-data ini dibagi lagi menjadi data pelatihan dan data pengujian di mana data pelatihan berisi 80% total data dan sisa data adalah data pengujian. Proses prediksi (*forecasting*) dilakukan selama 1440 menit dengan bantuan pembaharuan (*update*). Hasil dari proses prediksi memperlihatkan bahwa dengan melakukan pembaharuan secara berkala, hasil prediksi yang dihasilkan lebih baik dibandingkan hasil prediksi dengan tidak menggunakan pembaharuan. Pemilihan periode pembaharuan berdampak pada seberapa baiknya sistem saraf dapat memprediksi harga mata uang kripto.

Kata-kata kunci: *Bitcoin, Deret Waktu, Sistem Saraf Buatan, CNN, LSTM*

ABSTRACT

Profit is one of the many reasons that people start investing. Cryptocurrency is one of the many investment that people use these days. Forecasting the value of a certain cryptocurrency in the future will help investor take decisions for their investment. Neural networks can be used to forecast the value of cryptocurrency in the future. The neural networks or model that will be used to forecast are long short-term memory(LSTM) and a combination of long short-term memory and convolutional neural network. The error of the model is the benchmark for the quality of the model. The data consists of minutely-updated data from 2020 to 2021. It will be broken down into three parts. For every part, 80% of the data is the training data and 20% is the testing data. The forecasting process is done for 1440 minutes and assisted with updates. The results suggests that update have an impact on the forecasting process. The update period impacted the results of the forecasting.

Keywords: Bitcoin, Time Series, Neural Networks, Convolutional Neural Networks, Long Short-Term Memory

KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa karena atas rahmat-Nya yang melimpah penulis dapat menyelesaikan skripsi ini. Skripsi berjudul "Analisis Model *CNN* dan *LSTM* untuk Memprediksi Harga *Bitcoin*" disusun sebagai syarat untuk menyelesaikan studi Strata-1 di Program Studi Matematika, Fakultas Teknologi Informasi dan Sains (FTIS), Universitas Katolik Parahyangan (UNPAR), Bandung.

Terdapat hambatan dan kesulitan selama penyusunan skripsi ini. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada pihak yang telah membantu penulis selama penyusunan skripsi ini terutama kepada:

- Keluarga yang selalu memberikan dukungan, penghiburan selama proses penyusunan skripsi untuk penulis.
- Bapak Agus Sukmana, MSc selaku dosen pembimbing yang telah memberikan ilmu, arahan, dan saran dalam proses penyusunan skripsi.
- Bapak Dr. Andreas Parama Wijaya selaku dosen pembimbing yang telah memberikan ilmu, arahan dan saran dalam proses penyusunan skripsi.
- Ibu Maria Anestasia, M.Si, MAktSc selaku dosen penguji yang telah memberikan saran dan arahan untuk perbaikan skripsi.
- Bapak Taufik Limansyah, M.T. selaku dosen penguji yang telah memberikan saran dan arahan untuk perbaikan skripsi.
- Bapak Liem Chin, M.Si selaku koordinator skripsi yang telah memberikan arahan selama proses penyusunan skripsi ini.
- Bapak Janto Vincent Sulungbudi, S.Si. yang telah memberikan ilmu dan arahan dalam proses perkuliahan terutama mata kuliah *Deep Learning*.
- Seluruh dosen FTIS khususnya dosen Program Studi Matematika yang telah memberikan ilmu dan arahan selama perkuliahan.
- Seluruh staf Tata Usaha FTIS yang telah memberikan bantuan administrasi selama perkuliahan.
- Nasya Kanya dan Elbert Louis sebagai teman yang memberikan dukungan, semangat , dan bantuan selama perkuliahan dan proses penyusunan skripsi.
- Chihan Susanto sebagai teman yang selalu memberikan dukungan dan hiburan selama perkuliahan dan proses penyusunan skripsi.
- Teman-teman "Sukaipu" (Bryan, David, Wiksadana, Jason) sebagai teman-teman terdekat yang selalu memberikan dukungan dan hiburan selama proses perkuliahan dan penyusunan skripsi.
- Belinda Carensia sebagai teman yang selalu memberikan dukungan, semangat, dan hiburan selama perkuliahan dan proses penyusunan skripsi.
- Teman-teman kost *The Maple Place Residence* (Erick, Leo, Felix) yang selalu memberikan dukungan dan hiburan selama proses perkuliahan dan penyusunan skripsi.
- Teman-teman "JKT18" yang selalu memberikan dukungan dan hiburan selama proses perkuliahan dan penyusunan skripsi.
- Teman-teman divisi Design Kompetisi Matematika UNPAR, Buletin Tujuh Satu, dan program kerja lainnya yang tidak dapat disebutkan satu per satu yang telah memberikan ilmu dan pengalaman selama masa perkuliahan.

Penulis menyadari bahwa penulisan skripsi ini tidak luput dari adanya kekurangan. Oleh karena itu, penulis menerima segala kritik dan saran yang membangun dari penulis untuk penyempurnaan skripsi ini. Semoga skripsi ini dapat bermanfaat bagi pembaca.

Bandung, Januari 2022

Penulis

DAFTAR ISI

KATA PENGANTAR	xv
DAFTAR ISI	xvii
DAFTAR GAMBAR	xix
DAFTAR TABEL	xxi
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Batasan Masalah	2
1.5 Metodologi	2
1.6 Sistematika Pembahasan	2
2 LANDASAN TEORI	3
2.1 Deret Waktu	3
2.1.1 Stasioneritas [1]	3
2.1.2 Volatilitas	5
2.1.3 Pola Deret Waktu [1]	6
2.2 Pembelajaran Mesin	7
2.2.1 Komponen Sistem Saraf Buatan	7
2.2.2 Sistem Saraf Buatan Tersambung Penuh	13
2.2.3 Sistem Saraf Buatan Konvolusional	14
2.2.4 Sistem Saraf Berulang dan LSTM	16
2.2.5 <i>Backpropagation</i> dan <i>Forward Propagation</i>	18
2.2.6 Pelatihan Model	19
2.3 Mata Uang Kripto(<i>Cryptocurrency</i>)	19
2.3.1 <i>Blockchain</i>	21
2.3.2 Keamanan Mata Uang Kripto	21
2.3.3 Transaksi Mata Uang Kripto	21
3 METODOLOGI	23
3.1 Persiapan Data	23
3.1.1 Data	23
3.1.2 Persiapan Data	23
3.2 Pembangunan Model dan Penerapan	24
3.2.1 Pembangunan Model	24
3.2.2 Pelatihan Model	26
3.2.3 <i>Fitting</i> dan Prediksi Nilai Masa Depan	26
3.3 Evaluasi Model	28

4 HASIL DAN PEMBAHASAN	29
4.1 Hasil Pelatihan	29
4.1.1 Dataset 1	29
4.1.2 Dataset 2	30
4.1.3 Dataset 3	32
4.2 Hasil Fitting dengan Data Pengujian	34
4.2.1 Dataset 1	34
4.2.2 Dataset 2	35
4.2.3 Dataset 3	36
4.3 Hasil Prediksi tanpa/dengan Pembaharuan	38
4.3.1 Hasil Prediksi tanpa Pembaharuan	38
4.3.2 Hasil Prediksi dengan Pembaharuan	43
5 KESIMPULAN DAN SARAN	53
5.1 Kesimpulan	53
5.2 Saran	53
DAFTAR REFERENSI	55

DAFTAR GAMBAR

2.1	Contoh Deret Waktu Stasioner	4
2.2	Contoh Deret Waktu Tidak Stasioner	4
2.3	Harga Obligasi di Amerika Serikat	6
2.4	Produksi Listrik di Australia.	6
2.5	Penjualan Rumah di Amerika Serikat.	7
2.6	Sistem saraf buatan dengan 3 lapisan: masukan, tersembunyi, luaran.	8
2.7	Sistem saraf buatan antara lapisan masukan dan lapisan tersembunyi dengan ilustrasi bobot dan bias.	8
2.8	Ilustrasi Fungsi Aktivasi dari lapisan masukan ke lapisan tersembunyi.	9
2.9	Plot Fungsi Aktivasi Sigmoid.	10
2.10	Plot Fungsi Aktivasi Tanh.	10
2.11	Plot Fungsi Aktivasi ReLU.	11
2.12	Ilustrasi Fungsi Biaya dengan Titik Awal dengan Biaya yang Belum Minimum.	11
2.13	Ilustrasi Fungsi Biaya Mencapai Titik Optimal.	12
2.14	Ilustrasi Sistem Saraf Buatan Tersambung Penuh dengan 4 Lapisan.	14
2.15	Ilustrasi proses konvolusi pada lapisan konvolusi.	15
2.16	Ilustrasi <i>padding</i>	15
2.17	Ilustrasi Lapisan <i>Max Pooling</i>	16
2.18	Skema LSTM.	17
2.19	Perbesaran Sel pada LSTM.	17
2.20	Ilustrasi Sistem Saraf Buatan.	19
2.21	Ilustrasi <i>Ledger</i>	20
2.22	Ilustrasi <i>Genesis Block</i>	20
2.23	Ilustrasi <i>Blockchain</i>	21
3.1	Plot harga <i>bitcoin</i> pada tahun 2020-2021.	23
3.2	Skema Model LSTM 1.	25
3.3	Skema Model LSTM 2.	25
3.4	Skema Model Gabungan CNN - LSTM.	26
4.1	Hasil <i>Fitting</i> Data Pelatihan Model LSTM 1.	29
4.2	Hasil <i>Fitting</i> Data Pelatihan Model LSTM 2.	29
4.3	Hasil <i>Fitting</i> Data Pelatihan Model CNN-LSTM.	30
4.4	Hasil <i>Fitting</i> Data Pelatihan Model LSTM 1.	31
4.5	Hasil <i>Fitting</i> Data Pelatihan Model LSTM 2.	31
4.6	Hasil <i>Fitting</i> Data Pelatihan Model CNN-LSTM.	31
4.7	Hasil <i>Fitting</i> Data Pelatihan Model LSTM 1.	32
4.8	Hasil <i>Fitting</i> Data Pelatihan Model LSTM 2.	33
4.9	Hasil <i>Fitting</i> Data Pelatihan Model CNN-LSTM.	33
4.10	Hasil <i>Fitting</i> Data Pengujian Model LSTM 1.	34
4.11	Hasil <i>Fitting</i> Data Pengujian Model LSTM 2.	34
4.12	Hasil <i>Fitting</i> Data Pengujian Model CNN-LSTM.	35
4.13	Hasil <i>Fitting</i> Data Pengujian Model LSTM 1.	35

4.14 Hasil <i>Fitting</i> Data Pengujian Model LSTM 2	36
4.15 Hasil <i>Fitting</i> Data Pengujian Model CNN-LSTM.	36
4.16 Hasil <i>Fitting</i> Data Pengujian Model LSTM 1.	37
4.17 Hasil <i>Fitting</i> Data Pengujian Model LSTM 2.	37
4.18 Hasil <i>Fitting</i> Data Pengujian Model CNN-LSTM.	37
4.19 Hasil Prediksi tanpa Pembaharuan Model LSTM 1.	38
4.20 Hasil Prediksi tanpa Pembaharuan Model LSTM 2.	39
4.21 Hasil Prediksi tanpa Pembaharuan Model CNN-LSTM.	39
4.22 Hasil Prediksi tanpa Pembaharuan Model LSTM 1.	40
4.23 Hasil Prediksi tanpa Pembaharuan Model LSTM 2.	40
4.24 Hasil Prediksi tanpa Pembaharuan Model CNN-LSTM.	41
4.25 Hasil Prediksi tanpa Pembaharuan Model LSTM 1.	42
4.26 Hasil Prediksi tanpa Pembaharuan Model LSTM 2.	42
4.27 Hasil Prediksi tanpa Pembaharuan Model CNN-LSTM.	42
4.28 Hasil Prediksi dengan Pembaharuan Model LSTM 1 untuk Dataset 1.	44
4.29 Hasil Prediksi dengan Pembaharuan Model LSTM 2 untuk Dataset 1.	44
4.30 Hasil Prediksi dengan Pembaharuan Model CNN-LSTM untuk Dataset 1.	44
4.31 Hasil Prediksi dengan Pembaharuan Model LSTM 1 untuk Dataset 2.	46
4.32 Hasil Prediksi dengan Pembaharuan Model LSTM 2 untuk Dataset 2.	47
4.33 Hasil Prediksi dengan Pembaharuan Model CNN-LSTM untuk Dataset 2.	47
4.34 Hasil Prediksi dengan Pembaharuan Model LSTM 1 untuk Dataset 3.	49
4.35 Hasil Prediksi dengan Pembaharuan Model LSTM 2 untuk Dataset 3.	49
4.36 Hasil Prediksi dengan Pembaharuan Model CNN-LSTM untuk Dataset 3.	50

DAFTAR TABEL

3.1	Tabel data yang digunakan	24
3.2	Tabel Model yang Digunakan	26
4.1	Tabel Eror Data Pelatihan dengan Model untuk Dataset 1	30
4.2	Tabel Eror Data Pelatihan dengan Model untuk Dataset 2	32
4.3	Tabel Eror Data Pelatihan dengan Model untuk Dataset 3	33
4.4	Tabel Eror Data Pengujian dengan Model untuk Dataset 1	35
4.5	Tabel Eror Data Pengujian dengan Model untuk Dataset 2	36
4.6	Tabel Eror Data Pengujian dengan Model untuk Dataset 3	38
4.7	Tabel Eror Prediksi dengan Model untuk Dataset 1	39
4.8	Tabel Eror Prediksi dengan Model untuk Dataset 2	41
4.9	Tabel Eror Prediksi dengan Model untuk Dataset 3	43
4.10	Tabel Nilai MSE dengan Model untuk Dataset 1	45
4.11	Tabel Nilai RMSE dengan Model untuk Dataset 1	45
4.12	Tabel Nilai Eror Relatif dengan Model untuk Dataset 1	45
4.13	Tabel Nilai Koefisien Korelasi Nilai Real dan Hasil Prediksi untuk Dataset 1	46
4.14	Tabel Nilai MSE dengan Model untuk Dataset 2	48
4.15	Tabel Nilai RMSE dengan Model untuk Dataset 2	48
4.16	Tabel Nilai Eror Relatif dengan Model untuk Dataset 2	48
4.17	Tabel Nilai Koefisien Korelasi Nilai Real dan Hasil Prediksi untuk Dataset 2	48
4.18	Tabel Nilai MSE dengan Model untuk Dataset 3	50
4.19	Tabel Nilai RMSE dengan Model untuk Dataset 3	50
4.20	Tabel Nilai Eror Relatif dengan Model untuk Dataset 3	51
4.21	Tabel Nilai Koefisien Korelasi Nilai Real dan Hasil Prediksi untuk Dataset 3	51

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Salah satu alasan seseorang berinvestasi adalah untuk mendapatkan keuntungan. Salah satu alat investasi yang sedang terkenal saat ini adalah mata uang kripto. Mata uang kripto merupakan suatu alat tukar dan alat investasi yang ditemukan pada tahun 2009. *Bitcoin* merupakan mata uang kripto pertama yang ditemukan dan menjadi mata uang kripto yang sukses pada saat ini. Kelebihan dari mata uang kripto adalah transaksinya yang cepat dan bisa dilakukan dimanapun. Tetapi, dengan tidak adanya perantara, mata uang kripto juga dapat digunakan untuk hal-hal ilegal seperti pencucian uang.

Prediksi harga *Bitcoin* di masa yang akan datang dapat membantu untuk menentukan waktu pembelian dan penjualan *Bitcoin* sehingga investor dapat memaksimalkan keuntungan ataupun meminimalkan kerugian. Data harga *Bitcoin* di masa lalu dapat digunakan untuk memprediksi harga *Bitcoin* di masa yang akan datang. Data harga *Bitcoin* sepanjang periode waktu tertentu dapat dipandang sebagai suatu data deret waktu (*time series*) yang dapat digunakan untuk proses prediksi. Terdapat beberapa metode yang dapat digunakan untuk menganalisis suatu deret waktu, contohnya adalah *autoregression moving average* (ARMA) dan *autoregression integrated moving average* (ARIMA). Metode-metode tersebut menganalisis tren, pola musiman, dan juga pola siklik untuk memprediksi suatu harga di masa yang akan datang dalam suatu rentang waktu tertentu. Selain itu, terdapat metode lain yang berada pada lingkup kecerdasan buatan (*artificial intelligence*), yaitu dengan menggunakan pembelajaran mesin (*machine learning*) terutama dalam bidang saraf buatan (*artificial neural network*).

Saraf buatan terinspirasi dari cara kerja otak manusia. Layaknya sistem saraf pada manusia, sistem saraf buatan juga memiliki neuron. Jenis sistem saraf buatan yang akan digunakan di skripsi ini adalah LSTM dan CNN-LSTM. Kedua sistem saraf buatan tersebut mempunyai kelebihannya masing-masing. Model LSTM memiliki kemampuan untuk menganalisis hubungan antar data pada suatu rentang waktu tertentu. Model CNN-LSTM merupakan gabungan CNN dengan LSTM di mana CNN memiliki kemampuan untuk mengekstrak pola dari suatu data.

Teknik memprediksi harga *Bitcoin* dengan menggunakan pembelajaran mesin diawali dengan penciptaan suatu jenis sistem saraf buatan yang diberi nama LSTM (*Long Short-Term Memory*) [2] pada tahun 1997 oleh Sepp Hochreiter dan Jürgen Schmidhuber. Metode tersebut didesain untuk menganalisis data yang memiliki hubungan dalam suatu waktu yang panjang. Pada tahun 1999, diperkenalkan sistem saraf baru yang dikenal dengan CNN (*Convolutional Neural Network*) [3]. Tujuan awal dari pembuatan CNN adalah untuk pengenalan objek berbasis pembelajaran gradien di mana nilai minimum lokal dari fungsi konveks yang dicari dengan cara iteratif. Pada umumnya, analisis deret waktu hanya menggunakan LSTM, namun pada tahun 2020, penelitian yang dilakukan oleh Livieris mengungkapkan bahwa penggabungan model CNN dan LSTM dapat meningkatkan akurasi prediksi *Bitcoin* dibandingkan dengan hanya menggunakan LSTM saja [4].

Skripsi ini akan membahas prediksi harga *Bitcoin* dengan menggunakan sistem saraf buatan dan menganalisis jenis sistem saraf buatan yang baik untuk memprediksi harga *Bitcoin*. Metode yang akan dibahas pada skripsi ini adalah LSTM (*Long Short-Term Memory*) dan CNN-LSTM

(*Convolutional Neural Network-Long Short-Term Memory*).

1.2 Rumusan Masalah

Rumusan masalah yang akan dibahas pada skripsi ini adalah sebagai berikut:

1. Bagaimana cara penerapan model LSTM dan CNN-LSTM untuk memprediksi harga *Bitcoin*?
2. Bagaimana evaluasi hasil dari penerapan model LSTM dan CNN-LSTM untuk memprediksi harga *Bitcoin*?

1.3 Tujuan

Tujuan yang ingin dicapai dalam skripsi ini adalah:

1. Mengetahui pengaruh data pelatihan terhadap model LSTM dan CNN-LSTM.
2. Mengetahui kemampuan sistem saraf buatan dalam memprediksi deret waktu yang volatil.
3. Mengetahui perbandingan loss antara LSTM dan gabungan antara LSTM dan CNN dalam memprediksi *Bitcoin*.

1.4 Batasan Masalah

Batasan masalah pada skripsi ini adalah data harga *Bitcoin* di periode Januari 2020 - November 2021.

1.5 Metodologi

Metode yang digunakan dalam skripsi ini adalah pengambilan dan pengolahan data harga *Bitcoin*; pembuatan model dengan LSTM; dan CNN-LSTM. Pertama, nilai dari data harus diolah sehingga menjadi masukan untuk model yang dibuat. Data tersebut digunakan dalam model dengan LSTM dan CNN-LSTM. Terakhir, dilakukan evaluasi terhadap model yang sudah diterapkan. Bahasa pemrograman yang digunakan adalah bahasa *Python*.

1.6 Sistematika Pembahasan

Sistematika dalam penulisan skripsi adalah sebagai berikut:

1. Bab 1: Pendahuluan
Bab ini berisi tentang latar belakang, rumusan masalah, tujuan, batasan masalah, metodologi, dan sistematika pembahasan.
2. Bab 2: Landasan Teori
Bab ini berisi tentang teori yang mendukung skripsi ini. Materi yang dibahas meliputi deret waktu, pembelajaran mesin, dan mata uang kripto(*cryptocurrency*).
3. Bab 3: Metodologi
Bab ini berisi tentang langkah-langkah yang dilakukan dalam pengolahan data deret waktu, penerapan model dan evaluasi model. Model yang dibahas adalah LSTM, dan CNN-LSTM.
4. Bab 4: Hasil dan Pembahasan
Bab ini berisi tentang hasil dan pembahasan model yang digunakan untuk memprediksi harga *Bitcoin*.
5. Bab 5: Saran dan Kesimpulan
Bab ini berisi tentang saran dan kesimpulan dari skripsi ini.