

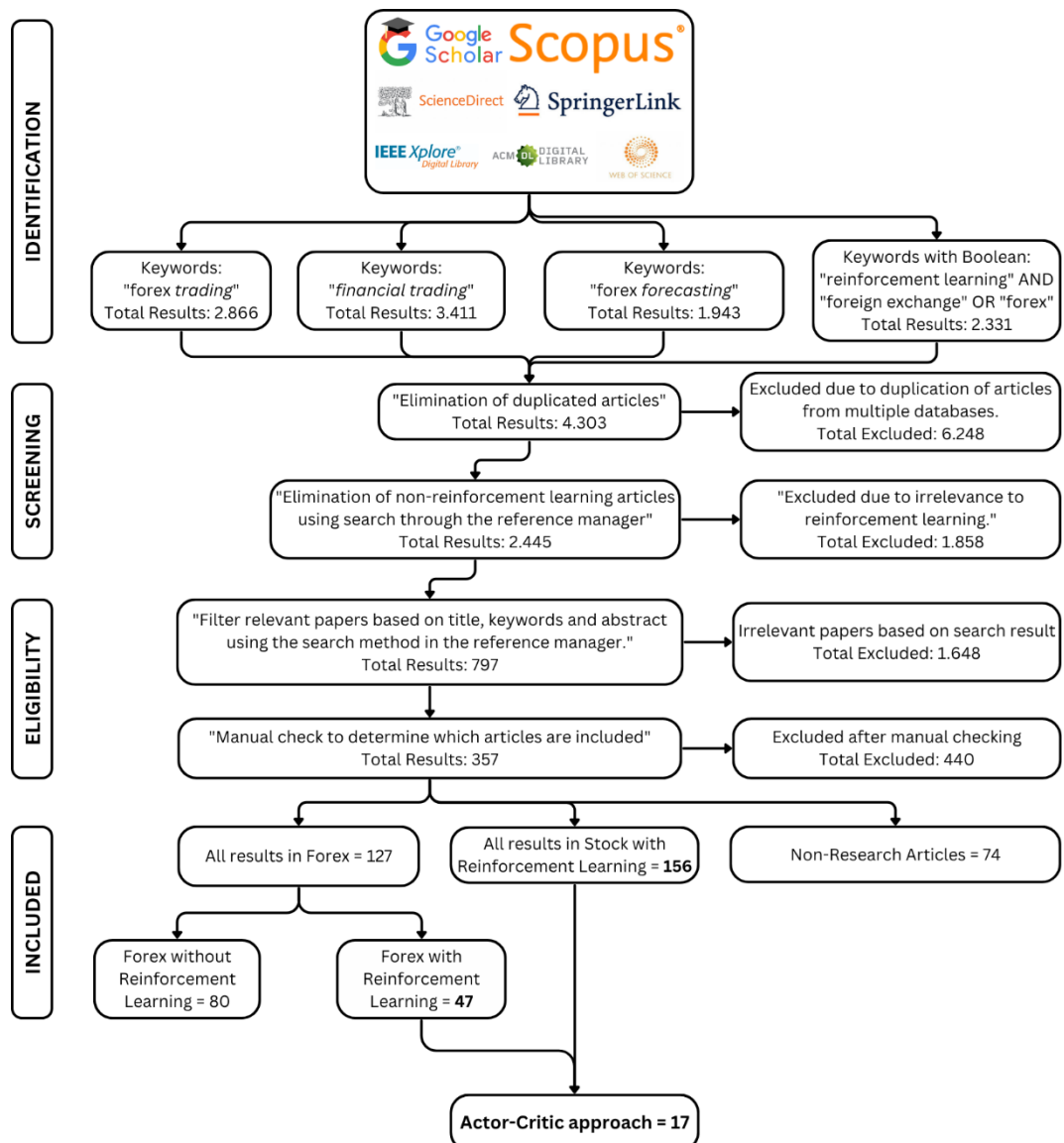
BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1. Tinjauan Pustaka

Kajian literatur dalam penelitian ini dilakukan melalui pendekatan *Systematic Literature Review* (SLR) untuk memperoleh pemahaman menyeluruh mengenai algoritma reinforcement learning yang telah diterapkan pada bidang perdagangan finansial, khususnya pasar valuta asing (*forex*). Proses penelusuran literatur dilakukan secara sistematis menggunakan pedoman SLR dari (Budgen & Brereton, 2006) serta prinsip *snowballing* dari (Wohlin, 2014). Penelusuran mencakup tujuh basis data utama yaitu, *Google Scholar*, *Scopus*, *SpringerLink*, *ScienceDirect*, *IEEE Xplore*, *ACM Digital Library*, dan *Web of Science* dengan rentang publikasi tahun 2015 hingga 2024. Pencarian dilakukan menggunakan kombinasi kata kunci “*forex trading*,” “*financial trading*,” “*forex forecasting*,” dan “*reinforcement learning*,” serta diolah menggunakan operator *Boolean* untuk memperluas cakupan hasil. Dari total 10.551 artikel yang diperoleh, dilakukan serangkaian penyaringan berlapis mencakup penghapusan duplikasi, pemeriksaan relevansi judul dan abstrak, serta penyisihan artikel *non-research*. Setelah melalui proses seleksi lima tahap, diperoleh 47 artikel utama yang secara spesifik membahas penerapan algoritma *reinforcement learning* dalam perdagangan valas.

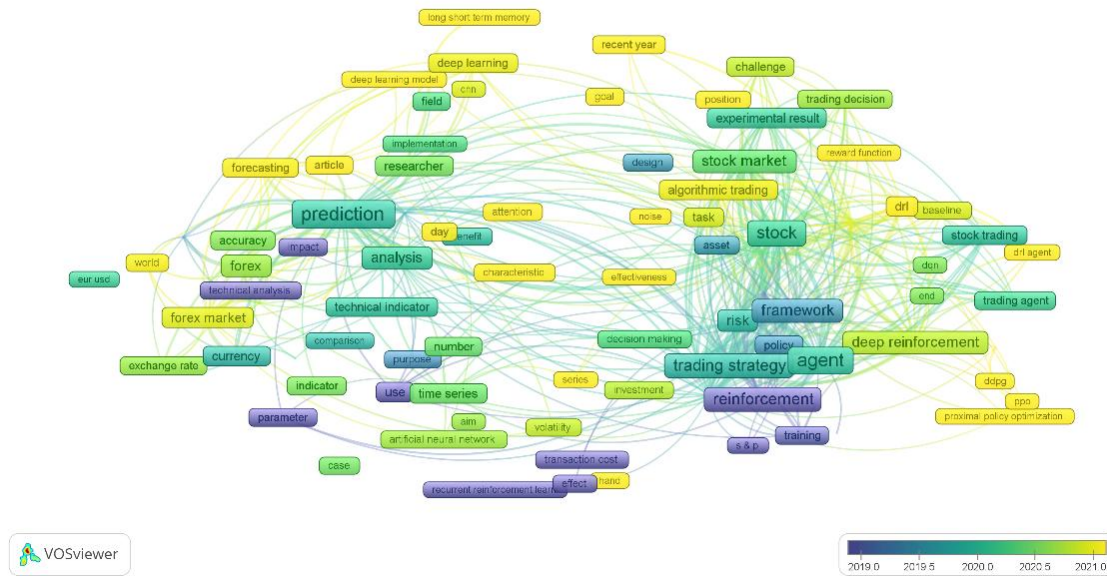
Melalui proses kajian literatur dari artikel-artikel yang telah dipilih pada tahap seleksi, ditemukan 17 penelitian bereputasi yang secara jelas menggunakan arsitektur *Actor–Critic* baik untuk memprediksi dinamika pasar (*financial forecasting*) maupun untuk merancang strategi perdagangan algoritmik. Temuan ini menjadi landasan utama pengembangan model dalam penelitian ini. Selain itu, melalui pendekatan analisis serupa, diperoleh pula 8 artikel penelitian utama yang membahas pemanfaatan pengumuman makroekonomi sebagai variabel dalam pengambilan keputusan perdagangan valas. Temuan ini menjadi dasar penting dalam merumuskan kerangka konseptual penelitian, karena menunjukkan keterkaitan yang kuat antara pengembangan model *reinforcement learning* berbasis *multi-critic* dan integrasi pengumuman makroekonomi.



Gambar 2.1 Diagram alir proses seleksi peninjauan artikel

Diagram alur diatas menggambarkan tahapan pencarian, penyaringan, dan seleksi akhir literatur hingga diperoleh kumpulan artikel yang layak untuk dianalisis lebih lanjut. Sementara itu, analisis berikutnya dilakukan melalui *VOSviewer* dengan metode *bibliographic coupling*, yang memetakan hubungan antar-publikasi berdasarkan referensi yang saling dikutip. Visualisasi jaringan ini menampilkan ukuran dan warna simpul (*nodes*) untuk menunjukkan jumlah referensi bersama, serta kekuatan keterkaitan antar-artikel yang menggambarkan kedekatan tematik dan tingkat relevansi di bidang *reinforcement learning for financial forecasting*.

Hasil pemetaan tersebut memberikan gambaran kuantitatif mengenai perkembangan penelitian, tren algoritma yang dominan, serta arah perkembangan riset ke depan di bidang ini.



Gambar 2.2 Analisis *bibliometrik* menggunakan *VOSViewer*

Temuan tersebut memperkuat arah penelitian selanjutnya yang berfokus pada pengembangan model *reinforcement learning* yang tidak hanya mampu memprediksi pergerakan harga, tetapi juga beradaptasi terhadap dinamika pasar secara waktu nyata. Sejalan dengan itu, penerapan *reinforcement learning* (RL) dalam perdagangan algoritmik telah menunjukkan potensi besar dalam menciptakan sistem pengambilan keputusan adaptif yang menyerupai pola pemrosesan kognitif manusia. Salah satu pendekatan utama dalam RL adalah algoritma *Actor-Critic*, yang mampu mempelajari kebijakan optimal berdasarkan imbalan jangka panjang (*reward*). Studi oleh (Deng et al., 2017) membuktikan efektivitas model *Deep Deterministic Policy Gradient* (DDPG) dan *Asynchronous Advantage Actor-Critic* (A3C) dalam mengoptimalkan manajemen portofolio di lingkungan pasar yang dinamis.

Model *Actor-Critic* umumnya diformulasikan dalam kerangka *Markov Decision Process* (MDP), seperti dalam penelitian (Tang, 2018), yang

memperhitungkan risiko dan biaya transaksi dalam strategi investasi multiperioda. Namun, pendekatan ini masih terbatas dalam konteks eksekusi pasar nyata, terutama ketika menghadapi delay, friksi, dan *slippage* harga. Untuk mengatasi tantangan tersebut, integrasi arsitektur *deep learning* seperti *Stacked Denoising Autoencoder* (SDAE) dan *Long Short-Term Memory* (LSTM) telah digunakan untuk memperkuat model DQN dan A3C (Li et al., 2019), serta adaptasi A3C pada indeks RTS yang mempertimbangkan biaya transaksi secara eksplisit (Ponomarev et al., 2019). Efisiensi dan stabilitas pembelajaran menjadi fokus pengembangan lebih lanjut melalui pendekatan multi-agen dan arsitektur nonlinier, seperti COSMOS berbasis *chaotic neuro-oscillatory learning* (R. S. T. T. Lee, 2019) dan GDQN berbasis GRU (X. Wu et al., 2020). Di sisi lain, (Yuan et al., 2020) mengembangkan *framework* DARL berbasis augmentasi data *candle* satu menit untuk meningkatkan ketahanan pelatihan terhadap statistik pasar seperti *skewness* dan *kurtosis*.

Salah satu tantangan utama dalam perdagangan algoritmik adalah kebutuhan untuk mengambil aksi dalam *continuous action space*. Algoritma seperti DDPG, TD3, dan *Soft Actor-Critic* (SAC) terbukti lebih fleksibel dibanding model diskrit (Nguyen & Luong, 2021), memungkinkan pengambilan keputusan yang lebih realistis dalam menentukan besaran posisi. Integrasi LSTM dalam POMDP oleh (Jia et al., 2021) menjadi landasan penting dalam menangani observasi parsial di pasar valas. Sementara itu, (H. K. Sagiraju & Mogalla, 2021) menguji ketahanan model terhadap kondisi ekstrem seperti pandemi COVID-19 melalui integrasi *multilayer perceptron* dan empat varian DRL. Pendekatan manajemen risiko juga terus dikembangkan melalui penerapan *Conditional Value-at-Risk* (CVaR) dalam struktur *Hierarchical* DDPG dan *Distributional* DDPG oleh (Wang & Ku, 2022).

Pengembangan model-model ini telah diperluas dengan memasukkan variabel-variabel baru. Sentimen pasar yang bersumber dari media sosial diintegrasikan oleh (K. Sagiraju & Mogalla, 2022) melalui *autoencoder* dan LSTM, sedangkan umpan balik posisi aktual digunakan oleh (Sun & Si, 2022) dalam kerangka SACRL-AF untuk menjembatani simulasi dan eksekusi riil. Untuk mendukung strategi *scalping* dalam perdagangan valas, (Tropmann-Frick et al., 2023) mengembangkan CNN

dengan *Gramian Angular Field*. (Yu et al., 2023) menggunakan *Nested RL* untuk adaptasi multilevel pada pasar valas yang sangat cepat berubah, sementara (Kong & So, 2023) menunjukkan bahwa pemilihan algoritma DRL yang sesuai dengan karakteristik pasar jauh lebih krusial dibanding hanya meningkatkan kompleksitas model.

Dalam konteks interpretabilitas sinyal dan pengurangan bias subjektif, pendekatan *fuzzy ensemble* DRL oleh (Hao et al., 2023) menyatukan lima Actor-Critic dalam model kolektif berbasis logika fuzzy. Studi oleh (S. Yang, 2023) memperkenalkan TC-MAC (mutual *Actor-Critic*) untuk memperkuat pemahaman konteks portofolio secara holistik. Terakhir, (Majidi et al., 2024) mengimplementasikan TD3 berbasis proporsi kas untuk pengambilan keputusan ukuran posisi, yang terbukti meningkatkan profitabilitas dan mengurangi eksposur pasar, sedangkan (Vishal et al., 2024) membangun *ensemble* DRL adaptif yang mampu memilih algoritma optimal berdasarkan *Sharpe Ratio*.

Salah satu rujukan penting dalam penelitian ini adalah pengembangan Multi-Critic DDPG oleh (J. Wu et al., 2018) yang menunjukkan bahwa penggunaan lebih dari satu *critic* dapat mengurangi sensitivitas terhadap kesalahan estimasi dan meningkatkan stabilitas pembelajaran dibanding DDPG standar. Namun, penelitian tersebut dikembangkan pada lingkungan kontrol terbatas, dengan definisi *state*, ruang aksi, dan fungsi *reward* yang berbeda jauh dari konteks perdagangan valuta asing. Penelitian ini mengadaptasi gagasan *multi-critic* tersebut ke dalam domain pasar valas yang sangat volatil dan dipengaruhi pengumuman makroekonomi terjadwal, dengan merancang arsitektur MC-DDPG yang menggunakan tiga *critic* independen untuk mengurangi *overestimation bias*, membangun *state* berbasis peristiwa yang menggabungkan data harga historis, indikator teknikal, dan puluhan jenis pengumuman makroekonomi. Dengan demikian, kontribusi penelitian ini terletak pada adaptasi dan pengembangan arsitektur MC-DDPG ke dalam lingkungan perdagangan valas berbasis pengumuman makroekonomi, yang belum secara spesifik dikaji dalam penelitian MC-DDPG sebelumnya.

Sejumlah penelitian untuk prediksi dan pengembangan strategi perdagangan, di mana indikator pengumuman makroekonomi digunakan sebagai input

fundamental yang disandingkan dengan data lain untuk meningkatkan akurasi. Penelitian oleh (Yıldırım et al., 2021), secara eksplisit menggabungkan data makroekonomi dengan indikator teknikal dalam model deep learning LSTM-nya, dan berhasil mencapai akurasi profit rata-rata 77-84%. Pendekatan serupa yang menggabungkan berbagai jenis data juga ditemukan pada riset (S. S. Lee & Wang, 2020) yang merancang strategi perdagangan tangguh dan mampu meningkatkan return tahunan sekitar 4,3%. Indikator makroekonomi yang umum digunakan dalam aliran penelitian ini meliputi Suku Bunga, Tingkat Inflasi, *Gross Domestic Product* (GDP), data ketenagakerjaan (seperti *Non-Farm Payrolls/NFP*), dan indeks saham. Selain data numerik, sentimen dari makroekonomi juga diolah, seperti yang dilakukan oleh (Coulombe, 2022), untuk menghasilkan model yang profitabilitasnya dapat diukur secara langsung.

Di sisi lain, terdapat aliran penelitian yang lebih menekankan pada analisis dampak dan pemahaman volatilitas pasar, yang menggunakan indikator makroekonomi sebagai pemicu utama untuk mengukur reaksi pasar. Lonjakan harga (*price jump*) dalam pasar keuangan umumnya didefinisikan sebagai perubahan harga yang sangat cepat dalam interval waktu yang pendek dan melebihi volatilitas normal. Studi oleh (Ben Omrane et al., 2020), secara spesifik menganalisis dampak dari rilis pengumuman makroekonomi dan menemukan bahwa efeknya terhadap volatilitas pasar dapat bertahan hingga dua jam. Penelitian ini mengamati beragam indikator fundamental berdampak tinggi dari AS dan Uni Eropa, seperti NFP, GDP, CPI, PPI, *Retail Sales*, dan ISM. Penelitian dalam kategori ini, termasuk studi peristiwa oleh (Kočenda & Moravcová, 2018), memberikan pemahaman mendalam tentang bagaimana pasar menyerap informasi dari indikator pengumuman makroekonomi tersebut.

Meskipun berbagai studi telah menunjukkan kemajuan signifikan dalam penerapan algoritma *deep reinforcement learning*, khususnya model *Actor-Critic* dalam konteks perdagangan algoritmik, masih terdapat sejumlah celah penting yang belum terjawab secara komprehensif, khususnya terkait dengan integrasi informasi fundamental dan teknikal dalam pengambilan keputusan otomatis. Oleh karena itu, penting dilakukan penelitian lanjutan yang mampu merespons dinamika pasar

secara lebih akurat dan adaptif. Berikut adalah ringkasan kesenjangan utama dan urgensi dilakukannya penelitian ini:

Kesenjangan Penelitian:

- Belum ditemukan model perdagangan valas yang mengimplementasikan arsitektur *Multi-Critic* DDPG secara spesifik untuk menangani masalah tingginya volatilitas harga valas akibat pengumuman makroekonomi.
- Informasi pengumuman makroekonomi terjadwal umumnya belum dijadikan fitur langsung dalam proses pengambilan keputusan model RL.
- Penggunaan indikator teknikal dan pengumuman makroekonomi masih dilakukan secara terpisah, tanpa mekanisme integratif yang mempertimbangkan keselarasan sinyal sebelum eksekusi transaksi.
- Penelitian sebelumnya belum mengevaluasi efektivitas model RL dalam merespons volatilitas pasar akibat deviasi nilai aktual terhadap pengumuman makroekonomi.

Pentingnya Penelitian:

- Pengumuman makroekonomi berdampak tinggi sering memicu lonjakan harga yang signifikan, sehingga dibutuhkan model yang mampu mempelajari dampak pengumuman makroekonomi terhadap perubahan pasar valas.
- Penggunaan tiga *critic* dalam MC-DDPG memberikan peluang untuk mengurangi bias estimasi (*overestimation*) nilai yang lazim terjadi pada model DDPG.
- Pendekatan ini berpotensi menghasilkan model prediksi harga valas yang lebih stabil dengan tingkat profitabilitas tinggi, khususnya dalam konteks pasar yang cepat berubah dan penuh volatilitas tinggi.

2.2. Keaslian (orisinalitas) Penelitian

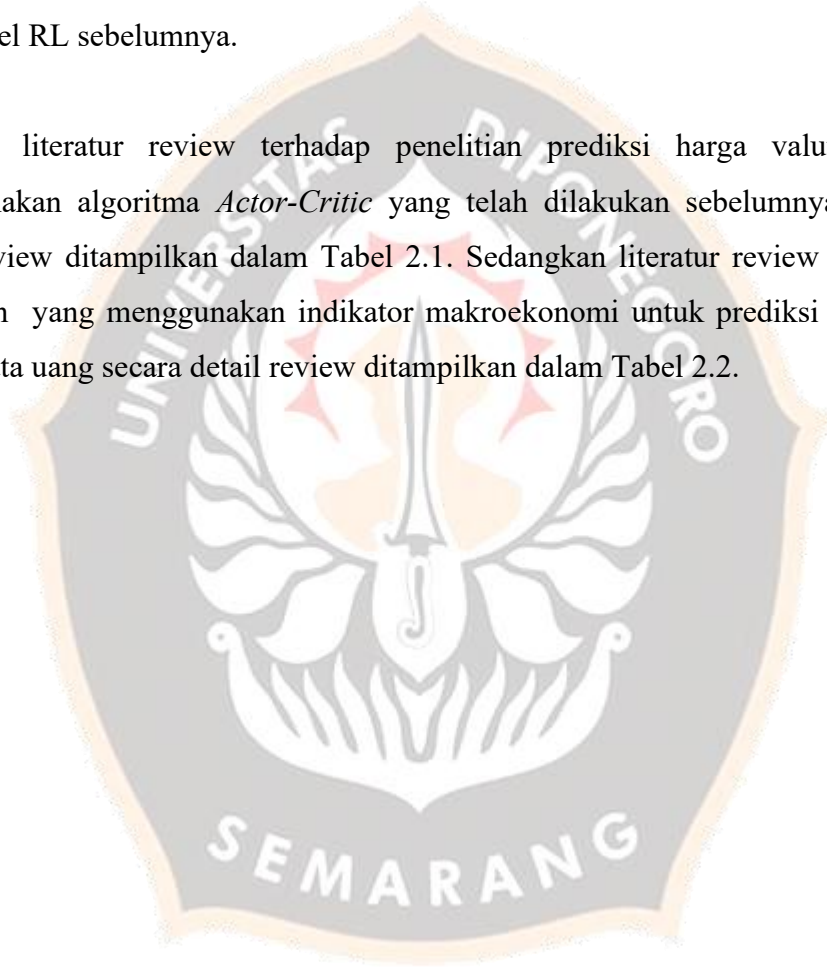
Penelitian ini memiliki keaslian pada pendekatan integratif dan arsitektur model yang diusulkan, yang secara spesifik belum ditemukan dalam studi sebelumnya. Keaslian tersebut dapat dijabarkan dalam poin-poin berikut:

- a. Penggunaan arsitektur Multi-Critic DDPG secara simultan untuk pasar valas, meskipun model DDPG dan TD3 telah banyak diterapkan, belum ditemukan penelitian yang secara eksplisit mengembangkan tiga *critic networks* independen untuk mengatasi prediksi harga valas dalam konteks pasar yang bereaksi terhadap pengumuman makroekonomi.
- b. Integrasi pengumuman makroekonomi secara waktu nyata sebagai fitur utama dalam model DRL, sebagian besar penelitian terdahulu hanya menggunakan data teknikal dan candle historis, atau memproses pengumuman makroekonomi sebagai data tambahan post-prediksi, bukan sebagai sinyal utama pengambilan aksi.
- c. Pemanfaatan data pengumuman makroekonomi dengan cakupan dan sifat yang unik, mencakup lebih dari 77 jenis dengan kategori antara lain:
 - Tenaga kerja & upah: *Non-Farm Payroll (NFP)*, *Unemployment Rate*, *Jobless Claims*, *Average Hourly Earnings*.
 - Inflasi & harga: *Consumer Price Index (CPI)*, *Core CPI*, *Producer Price Index (PPI)*.
 - Pertumbuhan ekonomi: *Gross Domestic Product (GDP)*, *Industrial Production*, *Retail Sales*.
 - Kebijakan moneter & suku bunga: *Interest Rate Decision (Federal Reserve, dll)*.
 - Indeks aktivitas & kepercayaan pasar: *Manufacturing PMI*, *Services PMI*, *Consumer Confidence Index*, *Business Confidence*, *IFO Business Climate*.
 - Perdagangan & neraca eksternal: *Trade Balance*, *Current Account Balance*, *Exports*, *Imports*.
- d. Selain itu penelitian ini menggunakan data berbasis peristiwa melalui kalender (*event-based*) dengan frekuensi tinggi mengikuti rilis resmi harian. Berbeda dengan studi sebelumnya yang hanya memakai berita ekonomi sebagai

sentimen positif atau negatif setelah rilis, pendekatan ini memungkinkan representasi makro yang komprehensif sekaligus penangkapan reaksi pasar secara langsung dan dinamis.

- e. Penggabungan sinyal indikator teknikal dan fundamental secara *decision-aligned*. Penelitian ini tidak hanya menggabungkan dua sumber data, tetapi juga mensyaratkan konsistensi arah antara sinyal teknikal dan fundamental sebagai syarat eksekusi transaksi, yang belum diterapkan secara eksplisit dalam model RL sebelumnya.

Dari literatur review terhadap penelitian prediksi harga valuta asing menggunakan algoritma *Actor-Critic* yang telah dilakukan sebelumnya, secara detail review ditampilkan dalam Tabel 2.1. Sedangkan literatur review terhadap penelitian yang menggunakan indikator makroekonomi untuk prediksi terhadap harga mata uang secara detail review ditampilkan dalam Tabel 2.2.



Tabel 2.1 Literature review algoritma *Actor-Critic* untuk prediksi keuangan

No.	Penulis	Market	Algoritma RL	<i>Actor</i> Model	<i>Critic</i> Model	Indicator	Dataset	Matriks Evaluasi	Hasil Prediksi
1	(Tang, 2018)	<i>Forex</i>	Actor-Critic Based	Policy Gradient	Q-learning	-	Historical Data	Average Return;	Average return is 2.08
2	(R. S. T. T. Lee, 2019)	<i>Forex</i>	Discrete-Time Actor-Critic RL Model (DTAC-RLM)	Gaussian Function Neural Network for Deep Reinforcement Learning (DRL)	Lee-oscillator technology	-	Historical Data	Maximum Drawdowns (MDD); Sharpe Ratio (SR); Average Return	Overall trading performance of 10.55% during the nine-month evaluation period
3	(Li et al., 2019)	<i>Stock</i>	Asynchronous Advantage Actor-Critic (A3C)	Advantage function	Deep Q-network (DQN)	MACD, MA, EMA, ATR, ROC.	Historical Data; Technical Indicator	Annualized Return (AR); Sharpe Ratio (SR)	Annualized return 59.4% Sharp ratio 1.7
4	(Ponomarev et al., 2019)	<i>Stock</i>	Asynchronous Advantage Actor Critic (A3C)	Advantage function	Artificial Neuron Network	-	Historical Data	Profit Ratio; Sharpe Ratio (SR)	Profit Ratio 66% per annum six months of historical data
5	(X. Wu et al., 2020)	<i>Stock</i>	Gated Deterministic Policy Gradient (GDPG)	Policy Gradient	Q-learning	MA, EMA, MACD	Historical Data; Technical Indicator	Profit Ratio; Sortino Ratio; Yield Curve Analysis	Sortino Ratio = 1.30
6	(Yuan et al., 2020)	<i>Stock</i>	Soft Actor-Critic (SAC)	Proximal Policy Optimization (PPO)	Deep Q-Learning	-	Historical Data	Annualized Return (AR); Sharpe Ratio (SR)	Annualized Return 12.98% Sharpe Ratio 6.69%

No.	Penulis	Market	Algoritma RL	Actor Model	Critic Model	Indicator	Dataset	Matriks Evaluasi	Hasil Prediksi
7	(H. K. Sagiraju & Mogalla, 2021)	Stock	Advantage Actor-Critic (A2C)	Proximal policy optimization (PPO)	Deep Q-networks (DQN)	-	Historical Data	Cumulative Return (CR); Annualized Return (AR); Maximum Drawdowns (MDD); Sharpe Ratio (SR); Sortino Ratio	Cumulative Return = 0.856%, Annual Return = 0.096%, Sharpe Ratio = 0.22, Sortino Ratio = 0.55, Drawdown = 100.267
8	(Nguyen & Luong, 2021)	Stock	Deep Deterministic Policy Gradient (DDPG)	Policy Gradient	Q-learning	-	Historical Data	Total Profit; Total Loss	-
9	(Jia et al., 2021)	Stock	LSTM-DDPG	Policy Gradient	LSTM	-	Historical Data	Return, Risk (DSR), Profit	Average Return: 10.55% (9-month period); DSR > Profit reward; Lower risk with higher consistency
10	(K. Sagiraju & Mogalla, 2022)	Stock	Advantage Actor Critic (A2C)	Policy Gradient	Deep Q-network (DQN)	Twitter Sentiment	Historical Data; News Sentiment	Cumulative Return (CR); Annualized Return (AR)	Annual return 13% and Cumulative returns of 64%
11	(Sun & Si, 2022)	Stock	Twin Delayed Deep Deterministic Policy Gradient (TD3)	Policy Gradient	Q-learning	-	Historical Data	Annualized Return (AR); Maximum Drawdowns (MDD); Sharpe Ratio (SR)	Annualized Return = 63.48%, Sharpe Ratio = 1.27, Max Drawdown = 32%
12	(Wang & Ku, 2022)	Stock	Hierarchical DDPG dan Distributional DDPG	Policy Gradient	Q-learning	-	Historical Data	CVaR; Risk-adjusted Return; Percentile-based Return	CVaR Reduction: 35%; $\hat{\mu}$ -Percentile Expected Return: 7.4%; Drawdown Risk Mitigated in COVID crash

No.	Penulis	Market	Algoritma RL	Actor Model	Critic Model	Indicator	Dataset	Matriks Evaluasi	Hasil Prediksi
13	(Tropmann-Frick et al., 2023)	Forex	Proximal Policy Optimization (PPO)	Gramian Angular Difference Field (GADF)	Generalized Advantage Estimation (GAE)	RSI, B&B, MFI, MACD.	Historical Data; Technical Indicator	Maximum Drawdowns (MDD); Total Profit; Total Loss	Max drawdown is 33.5%
14	(Vishal et al., 2024)	Stock	Ensemble DDPG, TD3, PPO dan SAC	Policy Gradient	Q-learning	MACD, CCI, ADX, RSI	Historical Data; Technical Indicator	Average Return; Maximum Drawdowns (MDD); Sharpe Ratio (SR)	Average Return: 2.08; Sharpe Ratio: 1.32; Max Drawdown: -8.75%
15	(Kong & So, 2023)	Stock	Advanced Actor Critic (A2C)	Policy Gradient	Q-learning	MACD, RSI, CCI, ADX	Historical Data; Technical Indicator	Cumulative Return (CR); Annualized Return (AR); Maximum Drawdowns (MDD); Sharpe Ratio (SR); Average Return; Annual Volatility	JPX30 CR 71.51%, AR 12.736%, AV 22.027%, SR 0.65, MDD 28.252%
16	(Hao et al., 2023)	Stock	Advantage Actor-Critic (A2C)	Policy Gradient	Deep Q-network	MACD, WR, RSI, CCI, ADX	Historical Data; Technical Indicator	Sharpe Ratio (SR); Total Return	Return rate 49% Sharpe ratio 0.627
17	(Majidi et al., 2024)	Stock	Twin-Delayed Deep Deterministic Policy Gradient (TD3)	Policy Gradient	Q-learning	MRMA & TFMA	Historical Data; Technical Indicator	Sharpe Ratio (SR); Total Return	Total Return 22.0%, Sharpe Ratio 0.80

Tabel 2.2 *Literature review* prediksi harga valas menggunakan indikator makroekonomi

No.	Authors	Proposed Method	Macroeconomic Indicators	Dataset	Evaluasi	Result
1	(Kočenda & Moravcová, 2018)	ESM <i>intraday</i> (1 menit) digunakan untuk mengukur abnormal return sebagai reaksi pasar terhadap pengumuman makroekonomi dan kebijakan moneter (ECB & Fed).	Zona Euro/Jerman (10): Ifo, ZEW, PMI Manufaktur & Jasa, Produksi Industri, GDP awal, Penjualan Ritel, Neraca Perdagangan, CPI Flash, PPI. AS (6+): NFP, PMI ISM, Produksi Industri, GDP awal, Penjualan Ritel, Neraca Perdagangan, Core Durable Goods, CPI, PPI. Moneter: Suku bunga & kebijakan ECB/Fed.	CZK/EUR, PLN/EUR, HUF/EUR, CZK/USD, PLN/USD, HUF/USD Periode: 3 Januari 2011 s/d 31 Desember 2015	Test untuk asimetri reaksi (good/bad/neutral news).	Ada peluang real-time short-term trading dari surprise news. Reaksi signifikan pada hampir semua kategori pengumuman makroekonomi. EUR pairs responsif ke makroekonomi Euro/Jerman (PMI, Ifo, ZEW, GDP), kuat pada <i>good news</i> . USD pairs responsif ke makroekonomi AS (NFP, GDP), kuat pada <i>bad news</i> . Reaksi terbesar: NFP & GDP (US, Germany).
2	(S. S. Lee & Wang, 2020)	GJR mengaitkan ukuran & frekuensi <i>currency jumps</i> dengan variabel ekonomi (intraday, harian, kuartalan), untuk memprediksi lonjakan ekstrem kurs dan merancang <i>jump-robust carry trade</i> .	Predictors: Intraday/Event-based: FOMC, NFP, GDP rilis, Neraca Perdagangan, Personal Income, PPI, CPI. Quarterly/Nasional: GDP, Suku bunga (forward premium), Pertumbuhan M1, Arus masuk FDI, Ekspor & Impor.	USD terhadap AUD, BRL, CAD, EUR, HUF, INR, JPY, KRW, NOK, NZD, PLN, RUB, SGD, ZAR, SEK, CHF, TRY, GBP (Januari 1999 s/d Desember 2015)	Regresi ukuran & intensitas jump vs variabel makro. Backtest carry trade dgn/tnp jump filter. Metrik: Return, Sharpe, skewness, certainty equivalent.	Pemicu utama jumps: Makroekonomi AS, terutama FOMC & NFP. GDP kuartalan: negatif dengan frekuensi jumps → ekonomi kuat = risiko lebih rendah. Carry trade: return 23–26% lebih rendah saat frekuensi jumps tinggi. Jump-robust carry trade: tingkatan return tahunan $\pm 4,3\%$ (+80%).
3	(Ben Omrane et al., 2020)	STR memodelkan dampak pengumuman makroekonomi pada return dan volatilitas EUR/USD secara nonlinier antar-rezim.	AS (High Impact): NFP & jobless claims; Unemployment; GDP (adv/prelim); Retail Sales; Housing Starts & Home Sales; ISM Mfg/Services; CPI & PPI; Consumer Confidence; Factory Orders; Trade Balance; FOMC (rate & minutes); Personal Spending; Business Inventories; Empire State & Philly Fed. UE (DE/FR/IT/ES): GDP; CPI; Industrial Production; Retail Sales; Business/Consumer	EUR/USD (5 menit interval). Periode: 1 November 2004 s.d 31 Maret 2014	STR estimation: deteksi efek nonlinier & state-dependent. Wald test: bandingkan ekspansi vs resesi. Polynomial decay: durasi dampak (≤ 2 jam pasca rilis). Seasonality adj.: ARCH harian, FFF, intraday.	US news lebih dominan ($\approx 2\times$ signifikan). Dampak US news: return 10–20 bp/SD. Dampak EU news: 7–12 bp. Volatilitas naik >85% kasus, durasi ≈ 2 jam pasca rilis.

No.	Authors	Proposed Method	Macroeconomic Indicators	Dataset	Evaluasi	Result
			Confidence; IFO; ZEW (Expectations/Current); Trade Balance; Unemployment.			
4	(Yıldırım et al., 2021)	ME-LSTM: faktor makroekonomi. TI-LSTM: indikator teknikal. Rule-based layer: gabungan untuk tingkatkan akurasi & eliminasi prediksi berisiko rendah.	Suku bunga: Jerman, Euro, Fed. Inflasi: Euro, AS. Indeks saham: S&P 500, DAX.	EUR/USD Periode: Januari 2013 s.d Januari 2018 (5 tahun). Data dibagi: 80% training, 20% testing	Eksperimen: 1, 3, 5-day ahead; Epochs: 50, 100, 150, 200. Metrik: <i>Profit Accuracy</i> = rasio prediksi arah benar ÷ total transaksi.	Baseline ME/TI-LSTM: akurasi rendah ($\pm 50-53\%$). Gabungan langsung: naik sedikit ($\pm 53\%$). Hybrid LSTM (rule-based): jauh lebih tinggi → 1-day: 77–78% (vs $\pm 52\%$). 3-day: 79% (kadang 100%, transaksi sedikit). 5-day: 84% (beberapa $>90\%$). Lebih konservatif (transaksi lebih sedikit) tapi lebih profitable, hindari prediksi berisiko.
5	(Wanniarachchi et al., 2021)	RF & SVM: klasifikasi arah tren. Linear Regression: ukur kekuatan fluktuasi. BPWDNN: hasilkan sinyal akhir.	CPI, NFP, PPI, Suku bunga, Pengangguran, GDP, Penjualan ritel.	Forex historis AUD/USD, EUR/CHF, USD/JPY (2000–2020, 20 tahun).	k-Fold Cross-Validation RMSE (Root Mean Squared Error) MSE (Mean Squared Error)	Profit Factor: 1.21 (6 bulan) 1.25 (1 tahun) Growth Balance: +4.88% (berdasarkan MetaTrader Strategy Tester).
6	(Djemo et al., 2021)	Ensemble ML (voting & bagging): gabungan Naïve Bayes, RF, QDA, SVM, dan K-NN.	Industrial Production (proxy income), M2. Interest Rate (I), Interest Rate Differential (IRD). CPI, Inflation Differential (INFD). Terms of Trade (TOT = P_x/P_m). NASDAQ, NIKKEI 225, FTSE 100, CAC40. Harga emas, harga minyak.	Kurs USD/ZAR, EUR/ZAR, GBP/ZAR, JPY/ZAR (Jan 1991 – Mei 2020, ± 30 thn, interval bulanan).	Accuracy, Precision, Recall, F-score, Cohen's Kappa (k), AUC (Area Under ROC Curve)	USD/ZAR: akurasi 67.6% GBP/ZAR: akurasi 70.4% JPY/ZAR: akurasi 69.0% EUR/ZAR: akurasi 69.2%
7	(Plihal, 2021)	Heterogeneous Autoregressive (HAR) Models dengan variabel dummy pengumuman makroekonomi terjadwal. Forecasting	(AS 72, Euro Area 57, UK 32, Jepang 26) Labour Market: NFP, ADP, Unemployment, Jobless Claims, JOLTS, Productivity, Labour Costs, dll.	Currency pairs: EUR/USD, GBP/USD, USD/JPY (Periode: 2009 - 2017 (9 tahun))	10-fold cross-validation (block method) untuk pilih di LASSO Loss functions: Mean Squared Error (MSE)	Model dengan makroekonomi (N-HAR-CJ) selalu outperform baseline HAR. Akurasi forecast volatilitas meningkat hingga 12.4% (EUR/USD); sekitar 9% untuk GBP/USD; 8% untuk USD/JPY Makroekonomi AS dominan terutama labour

No.	Authors	Proposed Method	Macroeconomic Indicators	Dataset	Evaluasi	Result
		volatilitas harian EUR/USD, GBP/USD, dan USD/JPY.	Housing & Construction: Building Permits, Housing Starts, Home Sales, Construction Spending, HPI, NAHB. Prices/Inflation: CPI, Core CPI, PPI, Import Price, PCE Core. Output & Growth: GDP QoQ, Industrial Production, Capacity Utilization, Orders (Durable/Factory/Cap Goods), Inventories, Consumption, Income, Spending. Confidence & Surveys: ISM (Mfg/Services/Prices), PMI, Consumer Confidence (Conference Board, Michigan), Fed Surveys (Empire, Philly, Dallas, Richmond), NFIB, Beige Book. Trade & External: Trade Balance, Current Account. Monetary Policy: FOMC Rate & Minutes, Fed Balance Sheet/Statements. Energy & Commodities: DOE/EIA inventories, NatGas, WASDE (corn, soybean, wheat), USDA planting/stocks, Vehicle Sales.	DukasCopy (forex quotes) & Bloomberg (macroeconomic announcements)	QLIKE (robust terhadap noise). Model Confidence Set (MCS test) untuk bandingkan performa model.	market (NFP, payrolls, unemployment), FOMC rate decision, energy inventories. GDP & trade balance relatif kurang penting.
8	(Neghab et al., 2024)	Kerangka ML interpretable untuk prediksi CAD/USD: Regresi terpenalti (LASSO, Ridge) Model pohon (ETR, XGBoost, LightGBM) Deep learning (GRU)	Harian: Oil (WTI), Gold (London fix), TSX, S&P 500, Eurodollar 3M LIBOR (ED). Bulanan: IR (interest rate parity), PPI parity, M1 (money supply), Unemployment, Industrial Production.	Periode studi 2009 - 2021. Target kurs CAD/USD (Bank of Canada, harian).	Metrik: NRMSE serta akurasi arah	Baseline (kalender): NRMSE $\approx$ 0.06. + Crude oil: error turun signifikan; contoh LGBM horizon 1-hari $\rightarrow$ NRMSE $\approx$ 0.04.

2.3. Landasan Teori

2.3.1. Valuta asing atau valas (*foreign exchange / forex*)

Dalam pasar valas, mata uang diperdagangkan dan dihargai secara berpasangan. Contohnya, pasangan mata uang EUR/USD dikuotasi 1,1256. EUR adalah mata uang dasar dan USD adalah mata uang pembanding. Order pada dasarnya adalah permintaan untuk melakukan transaksi untuk membuka posisi. Setelah transaksi dibuka, transaksi tersebut merupakan sebuah posisi. Mata uang dasar dapat dianggap sebagai posisi jual, yaitu menjual mata uang dasar dengan mata uang pembanding. Sebaliknya, mata uang pembanding adalah posisi beli pada pasangan mata uang. Dalam perdagangan valas, unit terkecil untuk memperdagangkan pasangan mata uang dikenal sebagai *lot* yang memiliki efek pengganda pada harga mata uang. Misalnya, jika seorang pedagang valas membeli satu *lot* standar EUR/USD akan mendapatkan keuntungan \$100 jika ia berhasil meraih 10 *pips* dengan menjual di harga 1,1266 setelah sebelumnya membeli di harga 1,1256. Ini karena pada satu *lot* standar, nilai setiap satu *pip* perubahan harga pada EUR/USD adalah \$10.

Berikut beberapa istilah dalam perdagangan valas.

a. Pasangan mata uang

Valas diperdagangkan dengan pergerakan atau performa satu mata uang terhadap mata uang lainnya. Biasanya diperdagangkan secara berpasangan dan berkorelasi satu sama lain, misalnya EUR/USD. Jika EUR/USD mengalami kenaikan nilai tukar, maka Euro semakin kuat terhadap Dollar AS, sedangkan penurunan nilai tukar menunjukkan Dollar AS semakin kuat terhadap Euro. Ada lebih dari 55 pasangan mata uang yang tersedia untuk diperdagangkan, termasuk EUR/USD, GBP/USD, AUD/USD, NZD/USD, USD/CAD, USD/CHF dan USD/JPY, yang merupakan pasangan mata uang utama yang menguasai lebih dari 85% dari total pasar valas (Rivero, 2020).

b. Leverage

Leverage dalam perdagangan valas mengacu pada peminjaman uang. *Leverage* memungkinkan pedagang valas untuk mengeksekusi posisi *lot* yang lebih besar dengan saldo yang sangat kecil (Lien, 2008). *Leverage* dinyatakan dalam rasio, misalnya 1:100, yang berarti seorang pedagang valas dapat mengendalikan posisi senilai 100 kali dari modal awalnya, sehingga dengan deposit \$1,000, seorang pedagang valas bisa mengendalikan posisi senilai \$100,000.

c. Harga Bid/Ask & Spread

Bid mengacu pada harga tertinggi yang harus dibayar pedagang untuk suatu aset, sementara harga *Ask* mengacu pada harga terendah yang akan diambil pedagang untuk aset yang sama. *Spread* adalah perbedaan harga antara permintaan dan penawaran (Maskey, 2021). Broker di pasar valas mendapatkan penghasilan melalui komisi untuk membuka posisi atau melalui *spread* antara harga. *Spread* dalam pasangan mata uang bervariasi berdasarkan volatilitas, likuiditas, dan volume pasangan mata uang.

Symbol	Bid	Ask
↘ EURUSD	1.06596	1.06606
↗ GBPUSD	1.23862	1.23869
↗ USDCHF	0.89752	0.89762
↗ USDJPY	147.708	147.720

Gambar 2.3 Harga beli (*bid*) / harga jual (*ask*)

d. Margin

Margin dalam perdagangan valas merujuk pada sejumlah uang yang diperlukan dalam akun pedagang valas sebagai jaminan untuk membuka posisi transaksi dengan *leverage*. *Margin* adalah uang jaminan yang disimpan di broker mereka untuk memungkinkan perdagangan dengan volume yang lebih besar melalui *leverage* (Murphy, 2015).

e. *Pips*

Pips adalah pergerakan terkecil dalam harga pasar valas, singkatan dari "*Percentage in Point*" (Adam Hayes, 2021). Pip adalah unit standar untuk mengukur perubahan harga, setara dengan 1/100 dari 1% atau angka desimal ke-4 (0.0001). Sebagai contoh, jika EUR/USD bergerak dari 1,1050 ke 1,1051 itu berarti pergerakannya adalah 1 *pip*. Demikian pula, jika GBP/JPY bergerak dari 139.50 ke 139.60, pergerakannya adalah 10 *pips*. (Kirkpatrick II & Dahlquist, 2010).

f. *Lot size*

Ukuran Lot dalam perdagangan valas adalah satuan standar untuk mengukur volume transaksi. Ukuran yang umum digunakan, yaitu *Lot Standar* (100.000 unit mata uang dasar), *Mini Lot* (10.000 unit), dan *Micro Lot* (1.000 unit). Pemilihan ukuran *lot* ini sangat krusial karena secara langsung menentukan tingkat risiko dan potensi keuntungan. (Financemagnates.com & Finance-Magnates, 2019).

g. *Stop Loss & Take Profit*

Dalam perdagangan valas, "*Stop Loss*" dan "*Take Profit*" adalah dua alat manajemen risiko penting. *Stop Loss* adalah perintah yang digunakan untuk menutup posisi secara otomatis ketika harga mencapai level tertentu, untuk membatasi kerugian. Contohnya, jika seorang pedagang valas membeli EUR/USD pada harga 1,2000 dan menetapkan *Stop Loss* pada 1,1950, posisi akan ditutup otomatis jika harga mencapai 1,1950 atau lebih rendah, membatasi kerugian pedagang valas. *Take Profit* di sisi lain, memungkinkan pedagang valas untuk mengunci keuntungan dengan menutup posisi ketika harga mencapai level keuntungan yang telah ditentukan sebelumnya. Menggunakan kedua alat ini membantu pedagang valas memiliki strategi keluar yang lebih disiplin, mengurangi pengaruh emosi dalam pengambilan keputusan, dan meningkatkan kemampuan untuk mengelola risiko serta mengambil keuntungan.

h. Volatility

Volatilitas secara umum mewakili variasi harga aset tertentu. Dengan kata lain, volatilitas mengukur seberapa besar dan seberapa cepat harga berubah dari waktu ke waktu. Volatilitas juga sering dilihat sebagai ketidakpastian atau ukuran risiko karena perubahan nilai aset yang cepat membuat aset tersebut kurang dapat diprediksi (Rosa, 2011). Oleh karena itu, potensi pengembalian di masa depan adalah kurang pasti dan manajemen aset dan risiko menjadi lebih mahal.

Salah satu faktor terpenting yang mempengaruhi volatilitas di pasar keuangan adalah berita-berita ekonomi. Pelaku pasar bereaksi terhadap informasi baru dan menyesuaikan ekspektasi mereka dan karenanya pedagang valas mengambil posisi yang sesuai. Oleh karena itu, ketika beberapa informasi baru muncul, biasanya akan meningkatkan jumlah transaksi, harga, dan juga volatilitas (Plíhal, 2021). Ada dua jenis pengumuman utama, terjadwal dan tidak terjadwal. Penelitian ini hanya berfokus pada pengumuman makroekonomi terjadwal karena tanggal rilisnya sudah diketahui sebelumnya sehingga cocok untuk dianalisis, tidak seperti berita ekonomi yang tidak terjadwal seperti pidato bank sentral yang tidak dijadwalkan atau berita geopolitik yang memengaruhi ekonomi.

i. Tick Data

Data *tick* dalam valas adalah informasi harga yang merekam perubahan terkecil dalam harga *bid* (harga jual) dan *ask* (harga beli) dari aset keuangan, seperti pasangan mata uang, dalam waktu nyata. Setiap *tick* merepresentasikan penyesuaian terkecil dalam harga dan memberikan detail yang sangat tinggi tentang aktivitas pasar, memungkinkan pedagang valas untuk menganalisis tren, volume perdagangan, dan pola harga dengan lebih mendalam guna mendukung pengambilan keputusan perdagangan yang lebih terinformasi (Johnson, 2010; O'hara, 1998).

Data *tick* terjadi sesuai dengan aktivitas pasar secara waktu nyata. Setiap kali terjadi perubahan kecil dalam harga *bid* atau *ask* pada aset keuangan, seperti pasangan mata uang dalam pasar valas, itu dianggap sebagai satu *tick*. Interval waktu antara setiap *tick* dapat bervariasi tergantung pada likuiditas dan aktivitas perdagangan pada pasangan mata uang tertentu. Pada saat pasar sangat aktif, bisa terjadi beberapa *tick* dalam hitungan detik atau bahkan dalam fraksi detik (Aldridge, 2013). Kecepatan terjadinya data *tick* dapat berfluktuasi seiring dengan volatilitas pasar dan aktivitas perdagangan yang berubah-ubah .

2.3.2. Analisa perdagangan valas

Ada dua pendekatan yang digunakan dalam melakukan analisis prediksi keuangan, yaitu analisis teknikal dan fundamental. Analisis keuangan berbasis data, analisis teknikal didasarkan pada harga pembukaan, harga tertinggi, harga terendah, harga penutupan dan volume pasar historis, sedangkan analisis fundamental didasarkan pada data yang dihasilkan dari berita ekonomi (Hu et al. 2015; Datta et al. 2021).

a. Analisa fundamental

Nilai tukar dari pasangan mata uang mencerminkan neraca perdagangan antara dua negara. Meskipun nilai fundamental sangat memengaruhi pergerakan harga, analisis fundamental tetap sulit digunakan untuk prediksi jangka pendek karena nilai fundamental sebenarnya tidak dapat diamati dengan pasti dan hanya direpresentasikan melalui indikator yang bersifat publik (Kaltwasser, 2010).

Tiga jenis data yang digunakan dalam menganalisis pasar valas dari perspektif analisis fundamental adalah, (1) besarnya jumlah permintaan dan penawaran valas, (2) banyaknya aktivitas pasar aset, dan (3) adanya peristiwa politik dan sosial (Korczak et al., 2016). Dalam konteks penelitian ini, pengumuman makroekonomi termasuk dalam kategori pertama karena data ekonomi memengaruhi ekspektasi pelaku pasar

terhadap nilai tukar melalui perubahan permintaan dan penawaran mata uang.

Ada beberapa pengumuman makroekonomi terjadwal yang dapat digunakan dalam perdagangan valas, diantaranya yaitu:

- *Balance Of Payment* (BOP) atau Neraca Pembayaran adalah pencatatan semua transaksi (impor, ekspor, investasi) antara suatu negara dan seluruh dunia selama periode tertentu periode tertentu misalnya, triwulanan, tahunan, dll. Neraca perdagangan positif terjadi ketika nilai total barang dan jasa yang dijual oleh produsen dalam negeri ke luar negeri melebihi nilai total barang dan jasa luar negeri yang dibeli oleh konsumen dalam negeri, dan sebaliknya (Nti et al., 2020).
- *Consumer Price Index* (CPI) atau Indeks Harga Konsumen untuk mengukur rata-rata tertimbang dari harga sejumlah barang dan jasa konsumen. Indeks ini merupakan ukuran yang sering digunakan untuk mengidentifikasi periode inflasi atau deflasi yang pada tingkat ekstrim dapat mempengaruhi nilai tukar (Giannellis & Koukouritakis, 2013).
- *Gross Domestic Bruto* (GDP) atau Produk Domestik Bruto adalah total nilai pasar dari semua barang jadi dan jasa yang diproduksi di dalam suatu negara dalam periode waktu tertentu. Ini adalah indikator untuk mengonfirmasi penilaian ekonomi tentang tren jangka panjang (Kadiri & Alabi, 2014). Dalam analisis fundamental, ini dapat digunakan untuk menilai dampak fluktuasi mata uang terhadap produksi dalam negeri

b. Analisa teknikal

Analisa teknikal adalah sekumpulan teknik untuk memeriksa data pasar historis dengan tujuan mengidentifikasi pola berulang yang dapat digunakan untuk memprediksi tren pasar di masa depan dalam jangka pendek (Picasso et al., 2019).

Analisis teknikal pertama kali diperkenalkan pada tahun 1800-an oleh Charles Dow (Murphy, 1999) untuk mempelajari inefisiensi pasar di pasar saham. Sejak awal 1970-an, setelah berakhirnya sistem perjanjian penilaian mata uang *Bretton Woods*, para pedagang telah menggunakan analisis teknikal untuk prediksi di pasar valas (Bordo, 2017). Analisis teknikal bertujuan untuk memprediksi perubahan masa depan di pasar keuangan dengan mencari pola berulang dalam data pasar historis, terutama harga dan volume, minat beli atau kontrak di pasar keuangan (Seth, 2021). Menggunakan perhitungan matematis yang diterapkan pada data historis untuk mengukur pergerakan harga dan secara empiris mengonfirmasi keberadaan pola yang berulang dan prediksi masa depan (Stasinakis & Sermpinis, 2014).

Dalam dunia perdagangan valas, para pedagang valas cenderung menggunakan indikator teknikal untuk memprediksi pergerakan harga di masa depan. Terdapat bukti keberhasilan yang lebih tinggi dalam menemukan pola berulang ketika indikator teknikal diterapkan pada cuplikan data yang diambil dalam frekuensi tinggi (Hongguang & Ping, 2015).

Berikut ini adalah indikator teknikal yang paling umum digunakan dalam analisa teknikal: RSI, CCI, SMA, MACD.

- ***Relative Strength Index (RSI)***

Relative Strength Index (RSI) adalah indikator *osilator momentum* yang dikembangkan oleh J. Welles Wilder pada tahun 1978. RSI digunakan untuk mengukur kecepatan dan perubahan pergerakan harga. Indikator ini dapat membantu mengidentifikasi kondisi *overbought* (dibelanjakan berlebihan) atau *oversold* (dijual berlebihan) dalam instrumen keuangan (Wilder, 1978).

RSI dihitung dengan formula berikut:

$$RSI = 100 - \left(\frac{100}{1+RS} \right) \quad (1)$$

RS (Relative Strength) = rata-rata kenaikan harga selama periode tertentu dibagi dengan rata-rata penurunan harga selama periode yang sama.

RSI bergerak dalam kisaran dari 0 hingga 100. Secara default, RSI sering dihitung dengan periode 14 hari.

Overbought dan *Oversold*: Secara umum, RSI di atas 70 dianggap *overbought*, dan RSI di bawah 30 dianggap *oversold*. Garis 50 pada RSI sering dianggap sebagai indikator momentum. Jika RSI berada di atas 50, momentum dianggap harga naik, dan jika di bawah 50, momentum dianggap harga menurun. Namun, batasan ini bisa disesuaikan tergantung pada kondisi pasar dan instrumen yang dianalisis.

- ***Simple Moving Average (SMA)***

Simple Moving Average (SMA) adalah salah satu indikator teknikal paling dasar dan sering digunakan. SMA adalah rata-rata aritmetika sederhana dari sejumlah data harga selama periode waktu tertentu. Indikator ini menghitung rata-rata harga penutupan untuk sejumlah periode waktu tertentu (Murphy, 1999).

SMA dihitung dengan formula berikut:

$$SMA = \frac{\sum_{i=1}^n \text{Harga Penutupan}}{n} \quad (2)$$

dengan n adalah jumlah periode yang dipilih.

SMA digunakan dengan kombinasi dua atau lebih SMA (dengan periode yang berbeda) untuk menghasilkan sinyal. Misalnya, *crossover* antara SMA 50-hari dan SMA 200-hari (*golden cross* dan *death cross*).

- **Exponential Moving Average (EMA)**

Exponential Moving Average (EMA) adalah salah satu jenis *moving average* yang memberikan bobot lebih pada harga-harga terbaru. Berbeda dengan SMA yang memberikan bobot yang sama untuk semua harga dalam periode tertentu, EMA memberikan bobot lebih pada data harga terbaru (Murphy, 2015). Ini menjadikan EMA lebih responsif terhadap perubahan harga dibandingkan dengan Simple Moving Average (SMA).

EMA dihitung dengan formula berikut:

$$\text{EMA hari ini} = \left(\frac{\text{Harga_Penutupan} - \text{EMA_kemarin}}{\text{span}} \right) + \text{EMA_kemarin} \quad (3)$$

- *span* adalah periode EMA. Faktor *smoothing* biasanya dihitung dengan $\frac{2}{\text{span}+1}$

Harga yang berada di atas EMA cenderung menunjukkan tren naik, sedangkan harga di bawah EMA cenderung menunjukkan tren turun. EMA dengan periode yang berbeda dapat digunakan bersama-sama untuk menghasilkan sinyal. Misalnya, *crossover* antara EMA 12-hari dan EMA 26-hari.

- **Average True Range (ATR)**

Average True Range (ATR) adalah indikator volatilitas yang menunjukkan tingkat harga aset. ATR dihitung berdasarkan rentang harga harian, termasuk perbedaan antara harga tertinggi dan terendah dan antara harga penutupan sebelumnya dan harga tertinggi atau terendah saat ini (Cohen, 2023).

Untuk menghitung *True Range* (TR) untuk setiap periode menggunakan rumus:

$$TR = \max \left(\begin{array}{l} (High - Low), \\ |High - Close_{previous}|, \\ |Low - Close_{previous}| \end{array} \right)$$

Kemudian, ATR dihitung sebagai rata-rata dari periode *True Range* n sebagai berikut:

$$ATR_{first} = \frac{\sum_{i=1}^n TR_i}{n}$$

Metode pemulusan (*smoothing*) eksponensial digunakan untuk periode berikutnya.

$$ATR_{current} = \frac{(ATR_{previous} \times (n - 1)) + TR_{current}}{n}$$

- ***Moving Average Convergence/Divergence (MACD)***

Moving Average Convergence/Divergence (MACD) adalah indikator tren yang menunjukkan hubungan antara dua *Moving Average Eksponensial* (EMA) dari harga penutupan sebuah sekuritas. MACD salah satu indikator teknikal yang paling populer dan sering digunakan untuk mengidentifikasi perubahan dalam kekuatan, arah, momentum, dan durasi dari tren harga pada bidang keuangan (Elder, 2014).

MACD dihitung dengan formula sebagai berikut:

- MACD Line: $EMA_{12hari} - EMA_{26hari}$
- Signal Line: EMA_{9hari} dari MACD Line
- Histogram MACD: MACD Line – Signal Line

Sinyal dasar dari MACD adalah ketika MACD Line melintasi Signal Line. Lintasan naik (*bullish crossover*) dapat menunjukkan potensi uptrend atau sinyal beli, sedangkan lintasan turun (*bearish crossover*) dapat menunjukkan potensi *downtrend* atau sinyal jual.

- **Commodity Channel Index (CCI)**

Commodity Channel Index (CCI) adalah indikator teknikal yang dikembangkan oleh Donald Lambert pada tahun 1980. Meskipun nama indikator ini mengandung kata "komoditas", CCI dapat digunakan untuk analisis berbagai jenis aset, termasuk saham, indeks, dan mata uang. CCI dirancang untuk mengidentifikasi siklus perubahan harga dan menentukan kondisi *overbought* atau *oversold* (Schlossberg, 2006).

CCI dihitung dengan formula berikut:

$$CCI = \frac{(TP - SMA(TP))}{MD \times D} \quad (4)$$

Di mana

- TP (*Typical Price*) = $\frac{(High + Low + Close)}{3}$
- $SMA(TP)$ adalah *Simple Moving Average* dari TP .
- MD adalah Deviasi Rata-rata dari TP .
- D adalah konstanta normalisasi (biasanya 0,015) yang menyesuaikan CCI agar sebagian besar berada dalam kisaran +100 hingga -100.

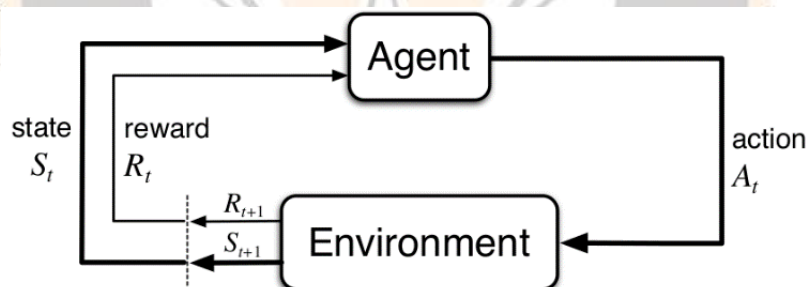
Zona Overbought & Oversold: Secara tradisional, pembacaan di atas +100 dianggap menunjukkan kondisi *overbought*, sementara pembacaan di bawah -100 dianggap menunjukkan kondisi *oversold*.

2.3.3. Reinforcement Learning

Reinforcement Learning (RL) merupakan salah satu cabang dari kecerdasan buatan yang beroperasi berdasarkan sistem hadiah (*reward*) dan hukuman (*punishment*) yang diterima dari umpan balik (*feedback*). Meskipun *Supervised Learning* (SL) juga memperhitungkan hadiah, SL umumnya hanya mempertimbangkan hadiah dari satu langkah aksi saja. Sebaliknya, RL bersifat berurutan dan agen RL yang ideal akan mempertimbangkan imbalan

akumulatif dalam jangka panjang. Model RL tidak memiliki label yang diawasi seperti dalam SL, namun mengandalkan umpan balik evaluatif yang diperoleh dari interaksi dengan lingkungan. Model ini menggunakan metode *trial and error* untuk mempelajari kebijakan yang optimal sepanjang waktu, dengan tujuan memaksimalkan akumulasi *reward* jangka panjang.

Pada setiap langkah waktu, model menerima pengamatan dari lingkungan dan memilih tindakan yang akan dieksekusi sesuai dengan kebijakan yang telah dipelajari. Lingkungan kemudian merespons dengan memberikan *reward* dan bertransisi ke keadaan baru sesuai dengan dinamika lingkungan. Dalam lingkungan episodik, interaksi ini berlanjut hingga model mencapai kondisi akhir dan memulai kembali prosesnya. Jika tidak ada keadaan akhir, lingkungan akan terus berlanjut. Selain itu, ada faktor diskon yang digunakan untuk mengukur pengaruh hadiah di masa depan terhadap keputusan saat ini. Model ini mengacu pada probabilitas transisi dan fungsi penghargaan (Sutton et al., 2018).



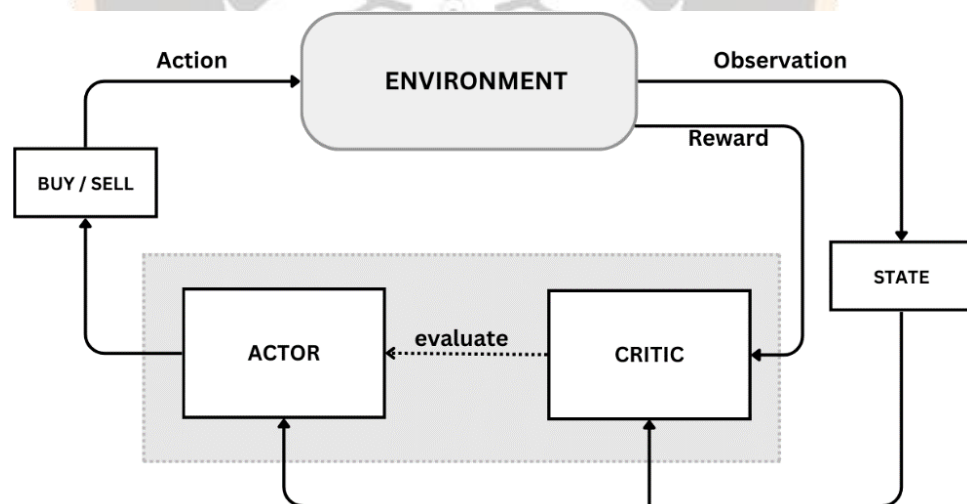
Gambar 2.4 Siklus kerja algoritma *reinforcement learning* (RL)

Dalam lingkungan non-linier, algoritma RL biasanya menggunakan aproksimator fungsi non-linier, seperti jaringan saraf, untuk mempelajari hubungan antara tindakan dan penghargaan. Fungsi aproksimator ini dapat memodelkan hubungan yang kompleks dan nonlinier antara input dan output model, memungkinkan algoritma RL untuk membuat keputusan yang efektif dalam lingkungan non-linier (Zhang & Mo, 2021). RL juga dapat mengoptimalkan sistem kontrol di industri manufaktur dan produksi energi. Dalam penggunaan umum, algoritma ini dapat mengoptimalkan arus lalu lintas

di jaringan transportasi, seperti jalan raya dan bandara. Selain itu, dapat digunakan untuk membuat keputusan perdagangan di bidang keuangan.

1. Actor-Critic Reinforcement Learning

Actor-Critic adalah algoritma yang berjalan secara paralel di beberapa model pelatihan *reinforcement learning* (RL) untuk meningkatkan efisiensi pelatihan. Algoritma ini menggabungkan kelebihan dari metode *policy gradient* (kebijakan) dan *value-function* (fungsi nilai) (Khuwaja et al., 2023). Dengan demikian, algoritma *actor-critic* menggunakan dua jaringan saraf yang terpisah: satu jaringan untuk mempelajari kebijakan (*policy network*) dan satu jaringan untuk mempelajari fungsi nilai (*value network*). Jaringan kebijakan memetakan kondisi (*state*) saat ini menjadi distribusi probabilitas atas tindakan yang mungkin dilakukan. Sementara itu, jaringan nilai memetakan kondisi-tindakan (*state-action*) menjadi perkiraan nilai jangka panjang yang diharapkan (*expected return*), yang kemudian digunakan sebagai umpan balik (*feedback*) untuk memperbaiki kebijakan.



Gambar 2.5 Diagram Actor-Critic reinforcement learning (RL)

Dalam kerangka *Actor-Critic*, agen menggunakan dua jaringan saraf: *actor* yang menentukan tindakan, dan *critic* yang mengevaluasi kualitas tindakan tersebut.

- *Actor*: adalah jaringan saraf yang menghasilkan distribusi probabilitas atas tindakan. Jaringan ini mengambil sebuah keadaan (atau pengamatan) dan memberikan kemungkinan untuk mengambil setiap tindakan yang mungkin dalam keadaan tersebut.
- *Critic*: Ini adalah jaringan saraf lain yang memperkirakan fungsi nilai baik nilai keadaan maupun nilai aksi. Jaringan ini akan menilai seberapa baik suatu keadaan atau tindakan dalam jangka panjang. Evaluasi *critic* ini menjadi dasar untuk memperbarui kebijakan *actor* agar semakin optimal.

Selama pelatihan, agen berinteraksi dengan lingkungan dengan memilih tindakan berdasarkan kebijakan dan menerima imbalan. Fungsi keuntungan (*advantage function*) digunakan untuk mengukur seberapa baik suatu tindakan dibandingkan dengan nilai rata-rata tindakan yang dapat diambil pada keadaan yang sama. Dengan demikian, fungsi ini menunjukkan keuntungan relatif dari suatu tindakan terhadap kebijakan saat ini.

2. *Advantage Actor-Critic (A2C)*

Advantage Actor-Critic (A2C) adalah algoritma *reinforcement learning* yang menggabungkan pendekatan *policy-based* dan *value-based* secara *synchronous*. A2C terdiri dari dua komponen utama: *actor* dan *critic*. *Actor* bertugas memilih aksi a berdasarkan kondisi *state* s , sedangkan *critic* mengevaluasi kualitas *state* tersebut melalui *state-value function* $V(s)$, yang digunakan untuk membentuk *advantage function* dan mengurangi *variance* pada estimasi gradien kebijakan.

a. *Advantage Function*

Advantage function $A(s_t, a_t)$ digunakan untuk menilai seberapa baik aksi tertentu dibandingkan dengan nilai ekspektasi keadaan *state* tersebut. Secara formal didefinisikan sebagai:

$$A(s_t, a_t) = Q(s_t, a_t) - V(s_t) \quad (5)$$

Di mana:

- s_t = state pada waktu ke- t
- a_t = aksi yang diambil pada *state* s_t
- $Q(s_t, a_t)$ = yaitu nilai ekspektasi return (reward kumulatif) jika aksi a_t dilakukan pada *state* s_t dan kemudian mengikuti *policy* saat ini. (Seberapa bagus kombinasi *state-action* ini ?)
- $V(s_t)$ = value dari *state* s_t , yaitu ekspektasi return dari keadaan s_t jika mengikuti *policy* saat ini. (Seberapa bagus keadaan ini secara umum, tanpa melihat aksi spesifik ?)

A2C memanfaatkan pengalaman terbaru untuk segera memperbarui *actor* dan *critic*. Aksi yang diambil dievaluasi dengan membandingkannya terhadap nilai rata-rata keadaan saat ini menggunakan fungsi keuntungan (*advantage*), sehingga kualitas aksi dapat diukur relatif terhadap ekspektasi kebijakan saat ini.

b. Keunggulan A2C

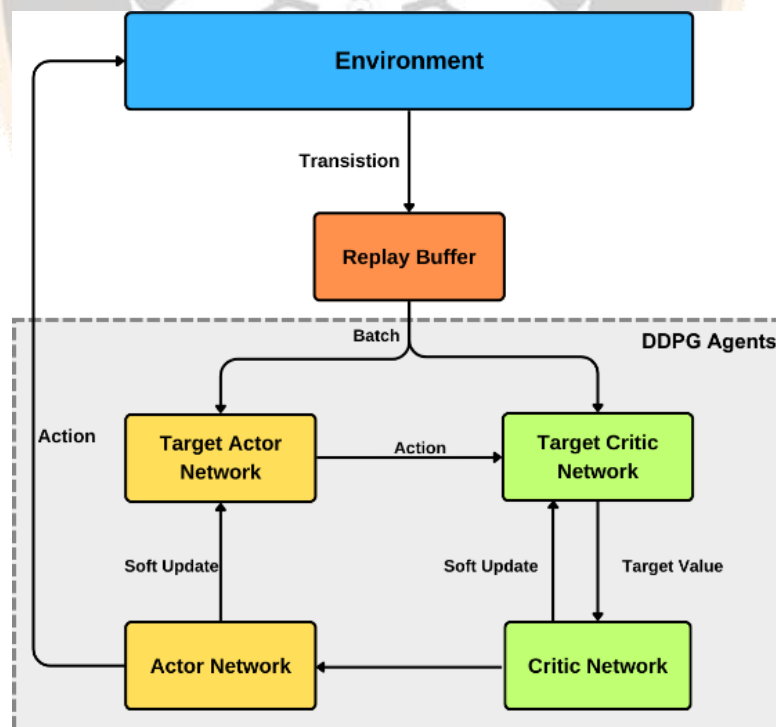
- Mengurangi variansi gradien dibanding metode *policy* gradien murni karena adanya *critic* yang berfungsi sebagai baseline.
- Dapat dijalankan secara paralel pada beberapa *environment* untuk mempercepat pelatihan.
- Relatif stabil pada masalah *discrete action space* dan struktur *reward* yang tidak terlalu noisy.

c. Kelemahan A2C

- Kurang fleksibel pada *continuous action space*, karena policy yang dihitung sebagai probabilitas diskret, tidak cocok untuk aksi bersifat kontinu.
- Sebagai algoritma *synchronous*, kadang *convergence* lebih lambat dibanding algoritma *off-policy* seperti DDPG atau SAC.
- Sensitif terhadap jumlah episode dan *reward shaping*, sehingga perlu *tuning hyperparameter* lebih cermat.

3. Deep Deterministic Policy Gradient (DDPG)

Deep Deterministic Policy Gradient (DDPG) adalah algoritma *reinforcement learning* yang berpengaruh yang menangani masalah dengan ruang aksi berkelanjutan (*continuous action space*) (Lillicrap et al., 2016). Belajar dari pengalaman masa lalu. DDPG menggunakan *Replay Buffer*, yang menyimpan ribuan pengalaman acak dari masa lalu. Saat belajar, ia mengambil *batch* acak dari buffer ini.



Gambar 2.6 Diagram model DDPG.

Proses dimulai ketika *Actor Network* memilih sebuah aksi untuk dilakukan di dalam *Environment*. Pengalaman dari interaksi ini, berupa (*state, action, reward, next state*), disimpan ke dalam sebuah memori yang disebut *Replay Buffer*. Untuk pembelajaran, agen mengambil sekumpulan data (*batch*) dari *Replay Buffer*. Proses belajar ini melibatkan empat jaringan: *Critic Network* utama belajar untuk mengevaluasi nilai dari sebuah aksi, sementara *Actor Network* utama belajar untuk memilih aksi yang terbaik. Untuk menjaga proses belajar tetap stabil, DDPG menggunakan *Target Actor* dan *Target Critic Network*, yang merupakan salinan dari jaringan utama yang diperbarui secara perlahan (*soft update*) untuk memberikan target pembelajaran yang tidak berubah terlalu cepat.

- a. *Target Network* (dihitung oleh *Target Critic*) Target ini digunakan untuk melatih *Critic Network* utama.

$$y = r + \gamma Q'(s', \mu'(s')) \quad (6)$$

Di mana:

- *y*: *Target Value*. Nilai yang diprediksi oleh *critic*. Jaringan *critic* akan dilatih untuk membuat prediksinya sedekat mungkin dengan nilai *y* ini.
- *r*: *Reward*. Imbalan yang diterima dari pengalaman yang diambil dari *replay buffer*.
- γ : *Gamma* (Faktor Diskon).
- *s'*: *Next State* (keadaan berikutnya) dari pengalaman di *replay buffer*.
- $\mu'(s')$: aksi target. Aksi yang akan dipilih oleh jaringan *Target Actor* (μ') di keadaan berikutnya.
- $Q'(s', \mu'(s'))$: *Target Q-value*. Prediksi dari jaringan target *critic* (Q') tentang nilai dari keadaan dan aksi (*state-action*) di masa depan.

- b. Pembaruan *Actor Network (Policy Gradient)*. Rumus ini menunjukkan bagaimana *Actor Network* belajar. *Actor* diperbarui ke arah yang disarankan oleh *Critic Network* untuk memaksimalkan nilai *Q-value*.

$$\nabla_a Q(s, \mu(s)) \mu(s) \quad (7)$$

Di mana:

- $\mu(s)$: Aksi saat ini. Aksi yang dihasilkan oleh *actor* utama (bukan target) untuk keadaan s .
 - $Q(s, \mu(s))$: Nilai *Q-value* yang diprediksi. *critic* utama mengevaluasi seberapa bagus aksi yang baru saja dipilih oleh *actor* utama.
 - $\nabla_a Q(s, \mu(s))$: Gradien terhadap aksi. Pembaruan *actor* dilakukan menggunakan *policy gradient deterministik*. sebagai arahan dari *critic* bagi *actor* untuk mengubah aksinya agar meningkatkan nilai *Q-value*.
- c. Keunggulan DDPG
- Mampu menangani *continuous action space*, yang tidak bisa dilakukan algoritma *policy gradient* diskret murni.
 - *Off-policy*, sehingga pengalaman yang tersimpan di *replay buffer* dapat digunakan berulang kali, meningkatkan efisiensi sample.
 - Menggunakan *target network* dan *replay buffer* untuk meningkatkan stabilitas training dan mengurangi korelasi sampel.
- d. Kelemahan DDPG
- Sensitif terhadap *noise* dan *hyperparameter* (misalnya *learning rate*, parameter untuk eksplorasi).
 - Rentan terhadap *overestimation bias* pada *Q-value*, yang dapat menyebabkan *divergence* jika *replay buffer* tidak memadai.
 - Lebih kompleks dibanding A2C dan membutuhkan lebih banyak tuning agar konvergensi yang stabil.

4. *Twin Delayed DDPG (TD3)*

TD3 merupakan pengembangan dari DDPG yang dirancang untuk mengatasi *overestimation bias* pada *Q-value* serta meningkatkan stabilitas pada *continuous action space* (Majidi et al., 2024; Park & Lee, 2021). TD3 menggunakan dua *critic network* (*twin critics*) dan *critic delayed policy update*, serta menambahkan *target policy smoothing* untuk mengurangi eksploitasi *noise* berlebihan terhadap *Q-value*.

a. TD3 memperkenalkan tiga inovasi utama:

1. *Clipped Double Q-Learning*: TD3 melatih dua jaringan *critic* (*twin networks*) dan menggunakan nilai *Q-value* minimum dari keduanya untuk menghitung target. Ini membantu mengurangi *overestimation bias*.
 2. *Delayed Policy Updates*: Jaringan Aktor dan *target networks* diperbarui lebih jarang daripada jaringan *critic*. Hal ini memungkinkan *critic* untuk memiliki estimasi nilai yang lebih akurat sebelum *actor* diperbarui.
 3. *Target Policy Smoothing*: Menambahkan sedikit *noise* pada aksi target dan melakukan *clipping* untuk membuat fungsi nilai menjadi lebih halus dan kebijakan lebih *robust* terhadap kesalahan estimasi.
- Dengan Target *Q-value* sebagai berikut:

$$y = r + \gamma \min_{i=1,2} Q'_i(s', \tilde{a}) \quad (8)$$

Di mana:

- y : *Target Value*. Sama seperti DDPG, ini adalah nilai yang harus dipelajari oleh *critic*.
- Q'_1, Q'_2 : (*Twin Networks*). TD3 menggunakan dua jaringan *critic* dan targetnya masing-masing.
- $\min=1,2$: Minimum dari dua nilai. Ini adalah trik utama TD3 dengan mengambil nilai *critic* minimum untuk mengatasi bias optimisme berlebih pada DDPG.
- $\tilde{a}$: Aksi *Target Actor* ($\mu'(s')$) yang diberi sedikit *noise* acak.

b. Keunggulan TD3

- Mengurangi *overestimation bias* pada nilai Q-value melalui penggunaan *twin critics* yang mengambil nilai minimum sebagai target.
- *Delayed policy update* membuat proses pembaruan lebih stabil selama training.
- *Target policy smoothing* meningkatkan ketahanan terhadap noise pada aksi dan estimasi nilai Q-value.
- Tetap cocok untuk *continuous action space*, dan lebih stabil dibanding DDPG.

c. Kelemahan TD3

- Lebih kompleks dibanding DDPG, membutuhkan dua *critic network* dan mekanisme *update policy* yang terpisah.
- Membutuhkan tuning tambahan pada parameter *smoothing* (variansi noise (σ) dan batas *clipping*).
- Menuntut kapasitas *replay buffer* yang besar dan training lebih berat secara komputasi dibanding algoritma sederhana seperti A2C.

5. Soft Actor-Critic (SAC)

SAC adalah algoritma *off-policy Actor-Critic* yang bertujuan memaksimalkan imbalan (*expected reward*), sekaligus memaksimalkan entropi kebijakan. Peningkatan entropi mendorong eksplorasi yang lebih luas, sehingga kebijakan menjadi lebih robus. SAC cocok untuk *continuous action space*, stabil terhadap *noise*, dan lebih efisien dibanding DDPG atau TD3 dalam beberapa kasus (Wen et al., 2021).

Inti algoritma dari SAC adalah, dimana suatu keadaan dianggap baik jika tidak hanya memberikan imbalan tinggi tetapi juga menawarkan potensi eksplorasi masa depan. Oleh karena itu, nilai sebuah aksi dihitung dari imbalan saat ini ditambah potensi imbalan dan eksplorasi di masa depan.

- a. SAC mengoptimalkan soft Q-value, yang memasukkan *reward* dan *entropy*:

$$Q(s_t, a_t) = r(s_t, a_t) + \gamma \mathbb{E}_{s_{t+1} \sim p}[V(s_{t+1})] \quad (9)$$

dengan *soft value function*:

$$V(s_t) = \mathbb{E}_{a_t \sim \pi}[Q(s_t, a_t) - \alpha \log \pi(a_t | s_t)] \quad (10)$$

Di mana

- s_t = *state* saat ini
- a_t = aksi *continue* yang dipilih oleh *policy* $\pi(a | s)$
- $r(s_t, a_t)$ = *reward* yang diterima setelah mengambil aksi a_t pada *state* s_t
- $\gamma \mathbb{E}_{s_{t+1} \sim p}[V(s_{t+1})]$ = Ekspektasi *Soft Value* di *state* berikutnya. Karena *soft value* sudah mengandung entropi, maka entropi juga berkontribusi pada nilai Q-value.
- $V(s_t)$ = *soft value function* yaitu nilai keadaan yang mempertimbangkan *reward* + entropi. Nilai satu keadaan $V(s_t)$ adalah ekspektasi (rata-rata) dari nilai Q dikurangi bonus *pinalti*.
- $Q(s_t, a_t)$: Nilai dari keadaan dengan mengambil aksi a_t di keadaan s_t sebelum mempertimbangkan bonus eksplorasi.
- α = *temperature parameter*, mengatur *trade-off* antara *reward* dan *entropy*
- $-\alpha \log \pi(a_t | s_t)$: Ini adalah bonus entropi. Bagian inilah yang membuat SAC "lunak" (*soft*). Ia memberikan nilai insentif pada keadaan di mana kebijakan (π) bisa bertindak lebih acak (memiliki entropi tinggi).

- b. Keunggulan SAC

- Memberikan eksplorasi stabil melalui *entropy regularization* sehingga kebijakan tidak cepat terkunci pada satu aksi.
- Lebih *robust* terhadap *noise* dibanding DDPG dan TD3.
- Cocok untuk *continuous action space* dan *reward* yang bersifat *stochastic*.
- *Off-policy*, sehingga *sample replay buffer* dapat digunakan efisien.

c. Kelemahan SAC

- Lebih kompleks secara implementasi dibanding A2C atau DDPG.
- Membutuhkan tuning parameter α untuk menjaga keseimbangan eksplorasi (entropi) dan eksploitasi (reward).
- Memiliki beban komputasi lebih tinggi karena *update policy* melibatkan perhitungan gradien entropi pada setiap iterasi.

DDPG, TD3, dan SAC merupakan keluarga algoritma *off-policy actor-critic* yang dirancang untuk menangani *continuous action space*, tetapi masing-masing mengadopsi strategi yang berbeda dalam menangani stabilitas pembelajaran, *overestimation bias*, dan eksplorasi. DDPG dapat dipandang sebagai algoritma dasar yang kemudian disempurnakan oleh TD3 melalui *twin critics* dan *delayed update*, sedangkan SAC menambahkan prinsip *maximum entropy reinforcement learning* untuk memperoleh kebijakan yang lebih eksploratif dan robust.

Dalam konteks penelitian ini, DDPG menjadi titik awal arsitektur karena kesesuaiannya dengan ruang aksi kontinu dan sifat *off-policy* yang efisien secara sampel. TD3 dan SAC memberikan inspirasi penting terkait penanganan *overestimation bias*, stabilitas, dan eksplorasi; gagasan-gagasan tersebut diadaptasi secara selektif ke dalam rancangan arsitektur *Multi-Critic* DDPG (MC-DDPG) yang dikembangkan pada penelitian ini.