

***Analisa Kinerja Infrastructure as a Service pada Cloud Computing
dengan Pendekatan Reinforcement Learning***

TUGAS AKHIR

**Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer**



Oleh:

Mutia Yasmin Azzahra

09011182126006

**JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA**

2024

HALAMAN PENGESAHAN

Analisa Kinerja *Infrastructure as a Service* pada *Cloud Computing* dengan Pendekatan *Reinforcement Learning*

SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

Mutia Yasmin Azzahra

09011182126006

Palcbang, 8 Januari 2025

Mengetahui,

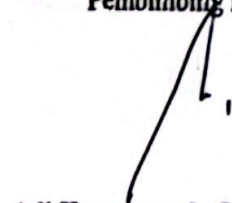
Pembimbing I



Dr. Ahmad Hervanto, M.T.

NIP. 198701222015041002

Pembimbing II



Adi Hermansyah, S.Kom., M.T.

NIP. 198904302024211001

Ketua Jurusan Sistem Komputer



Dr. Ir. H. Sukemi, M.T.

NIP. 19961203200604100

AUTHENTICATION PAGE

***Performance Analysis of Infrastructure as a Service in Cloud Computing
Using a Reinforcement Learning Approach***

SKRIPSI

***Submitted To Complete One Of The Requirements For
Obtaining A Bachelor's Degree in Computer Science***

By:

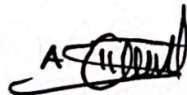
Mutia Yasmin Azzahra

09011182126006

Palembang, 8 January 2025

Acknowledge,

Final Project Advisor I



Dr. Ahmad Hervanto, M.T.

NIP. 198701222015041002

Final Project Advisor II



Adi Hermansyah, S.Kom., M.T.

NIP. 198904302024211001

Head Of Computer System Departement



Dr. Ir. H. Sukemi, M.T.

NIP. 19961203200604100

iii

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada:

Hari : Senin

Tanggal : 30 Desember 2024

Tim Penguji:

1. Ketua : Huda Ubaya, M. T.

2. Peguji : Prof. Deris Stiawan, M. T., Ph. D.

3. Pembimbing I : Dr. Ahmad Heryanto, M.T.

4. Pembimbing II : Adi Hermansyah, S.Kom., M.T.




Mengetahui, *2/1/25*
Ketua Jurusan Sistem Komputer



Dr. Ir. H. Sukemi, M.T.

NIP. 19961203200604100

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Mutia Yasmin Azzahra

NIM : 09011182126006

Judul : *Analisa Kinerja Infrastructure as a Service pada Cloud Computing dengan Pendekatan Reinforcement Learning*

Hasil Pengecakan Plagiat / Turnitin : 11%

Menyatakan bahwa laporan tugas akhir ini adalah hasil karya saya sendiri dan tidak mengandung unsur penjiplakan atau plagiat. Saya sepenuhnya menyadari bahwa jika terbukti adanya penjiplakan atau plagiat dalam laporan tugas akhir ini, saya siap menerima sanksi akademik dari Universitas Sriwijaya. Pernyataan ini saya buat dengan kesadaran penuh dan tanpa adanya unsur paksaan dari pihak manapun.



Palembang, 01 Januari 2025

Yang menyatakan



Mutia Yasmin Azzahra

NIM 09011182126006

Analisa Kinerja *Infrastructure as a Service* pada *Cloud Computing* dengan Pendekatan *Reinforcement Learning*

Mutia Yasmin Azzahra

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

Email:

ABSTRAK

Cloud Computing, khususnya layanan *Infrastructure as a Service* (IaaS), memberikan fleksibilitas tinggi, efisiensi biaya, dan skalabilitas dalam pengelolaan sumber daya komputasi. Namun, fluktuasi permintaan dan dinamika beban kerja memerlukan strategi manajemen yang adaptif untuk menjaga kualitas layanan. Penelitian ini mengeksplorasi penggunaan algoritma *Reinforcement Learning* (RL), yaitu *Deep Q-Network* (DQN) dan *Proximal Policy Optimization* (PPO), untuk mengoptimalkan kinerja IaaS. Pada percobaan yang menggunakan dataset ieeedatapot *metrics-with-output*, agen RL berhasil mempelajari pola penggunaan sumber daya cloud secara dinamis. Hasil penelitian menunjukkan bahwa DQN mencapai akurasi 89% dalam memprediksi status sistem, sementara PPO mencapai akurasi 82%. Hasil menunjukkan DQN memiliki performa sedikit lebih baik, dengan batch size 32 dibandingkan PPO yang membutuhkan 64. Kedua model menggunakan arsitektur jaringan identik [64, 64] dan learning rate serupa (DQN: 0.00025, PPO: 0.0003). Selain itu, kedua algoritma mampu mengurangi pemborosan sumber daya dengan meningkatkan efisiensi pemanfaatan CPU, memori, bandwidth, dan waktu respons. Penelitian ini memberikan kontribusi dalam optimalisasi manajemen IaaS menggunakan RL, dengan peluang untuk pengembangan lebih lanjut melalui integrasi algoritma lain dan penerapan pada skenario cloud computing yang lebih luas.

Kata Kunci: *Cloud Computing, Infrastructure as a Service* (IaaS), *Reinforcement Learning, Deep Q-Network* (DQN), *Proximal Policy Optimization* (PPO)

Mengetahui,

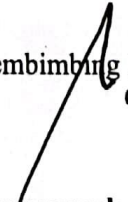
Pembimbing I



Dr. Ahmad Hervanto, M.T.

NIP. 198701222015041002

Pembimbing II



Adi Hermansyah, S.Kom., M.T.

NIP. 198904302024211001

Performance Analysis of Infrastructure as a Service in Cloud Computing Using a Reinforcement Learning Approach

Mutia Yasmin Azzahra

*Departement of Computer, Faculty of Computer Science, Sriwijaya
University*

Email:

ABSTRACT

Cloud Computing, particularly Infrastructure as a Service (IaaS), offers high flexibility, cost efficiency, and scalability in managing computing resources. However, fluctuating demands and dynamic workloads require adaptive management strategies to maintain service quality. This study explores the use of Reinforcement Learning (RL) algorithms, specifically Deep Q-Network (DQN) and Proximal Policy Optimization (PPO), to optimize IaaS performance. Using simulations based on datasets from IEEE DataPort, RL agents were designed to dynamically learn resource usage patterns. The results show that DQN achieved 89% accuracy in predicting system status, while PPO achieved 82%. Results indicate DQN performs slightly better, with a batch size of 32 compared to PPO's 64. Both models utilize identical network architectures [64, 64] and similar learning rates (DQN: 0.00025, PPO: 0.0003). Additionally, both algorithms reduced resource wastage by improving the efficiency of CPU, memory, bandwidth, and response time usage. However, challenges remain in addressing imbalances in negative class detection. This research contributes to the optimization of IaaS management using RL, with potential for further development through the integration of other algorithms and application to more complex cloud computing scenarios.

Keyword: *Cloud Computing, Infrastructure as a Service (IaaS), Reinforcement Learning, Deep Q-Network (DQN), Proximal Policy Optimization (PPO)*

Acknowledge,

Final Project Advisor I



Dr. Ahmad Heryanto, M.T.

NIP. 198701222015041002

Final Project Advisor II



Adi Hermansyah, S.Kom., M.T.

NIP. 198904302024211001

KATA PENGANTAR

Dengan penuh rasa syukur dan kerendahan hati, saya panjatkan puji dan syukur ke hadirat Allah SWT yang telah melimpahkan rahmat, karunia, dan kekuatan-Nya sehingga saya dapat menyelesaikan tugas akhir ini. Tanpa petunjuk dan kasih sayang-Nya, perjalanan panjang yang penuh tantangan ini tidak mungkin dapat saya lalui.

Persembahan pertama ini saya tujukan kepada kedua orang tua tercinta dan keluarga besar yang telah menjadi pilar utama dalam hidup saya. Doa-doa yang mereka panjatkan tanpa lelah, cinta tanpa syarat, dan dukungan yang tiada henti menjadi pelita yang menerangi jalan saya. Segala usaha dan pencapaian ini adalah bentuk kecil dari rasa terima kasih saya kepada mereka, meskipun tidak akan pernah sebanding dengan pengorbanan mereka.

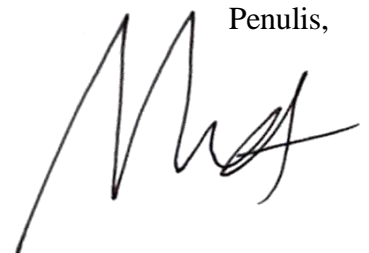
1. Ucapan terima kasih yang tulus juga saya sampaikan kepada Prof. Dr. Erwin, S.Si., M.Si., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya, atas bimbingan dan kebijaksanaannya yang menginspirasi kami sebagai mahasiswa untuk terus belajar dan berkembang.
2. Tak lupa, penghargaan mendalam saya sampaikan kepada Dr. Ir. Sukemi, M.T., selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya, yang telah memberikan arahan serta dukungan penuh selama proses studi saya.
3. Kepada kedua pembimbing saya, Bapak Dr. Ahamd Heryanto, M.T., dan Bapak Adi Hermansyah, S.Kom., M.T., terima kasih atas kesabaran, ilmu, dan waktu yang telah dicurahkan untuk membimbing saya dalam pengerjaan tugas akhir ini. Arahan dan masukan Bapak menjadi pondasi kokoh dalam menyelesaikan penelitian ini.
4. Saya juga mengucapkan terima kasih kepada dosen pembimbing akademik saya, Bapak Ahmad Fali Oklilas, M.T., yang dengan penuh perhatian memberikan nasihat dan motivasi selama masa studi saya di Universitas Sriwijaya.
5. Tidak lupa pula, ucapan terima kasih saya haturkan kepada Ibu Sari Nurhaliza, yang selalu sigap membantu administrasi akademik dengan penuh dedikasi dan kesabaran, mempermudah perjalanan saya selama ini.

6. Saya juga ingin mengungkapkan rasa terima kasih kepada teman-teman di berbagai grup WhatsApp yang menjadi tempat saya berbagi canda, keluh kesah, serta semangat selama proses pengerjaan tugas akhir ini. Terima kasih kepada teman-teman di grup Cubitus, grup Hayolo, grup Kalidoni & 3 Iilir, grup Bagi Pasangan, hingga grup Proyek TA Pak Adi. Dukungan kalian, baik berupa kata-kata, ide, maupun motivasi, begitu berarti bagi saya.
7. Tak lupa, penghargaan ini saya persembahkan kepada diri saya sendiri. Terima kasih telah bertahan, berjuang, dan tidak menyerah di tengah segala rintangan. Perjalanan ini tidak mudah, tetapi saya bangga pada diri saya yang tetap kuat dan konsisten.
8. Akhir kata, kepada siapa saja yang tidak dapat saya sebutkan satu per satu namun telah berkontribusi, baik secara langsung maupun tidak langsung, saya ucapkan terima kasih yang sebesar-besarnya. Semoga segala kebaikan dan dukungan yang telah diberikan mendapatkan balasan yang terbaik dari Allah SWT.

Penulis menyadari bahwa laporan ini masih sangat jauh dari kata sempurna. Untuk itu kritik dan saran yang membangun sangatlah diharapkan penulis. Akhir kata penulis berharap, semoga proposal tugas akhir ini bermanfaat dan berguna bagi rekan lainnya.

Palembang, 07 Januari 2025

Penulis,



Mutia Yasmin Azzahra

NIM 09011182126006

DAFTAR ISI

HALAMAN PENGESAHAN	ii
<i>AUTHENTICATION PAGE</i>	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PERNYATAAN.....	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL.....	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Perumusan dan Batasan Masalah.....	5
1.2.1 Batasan Masalah	5
1.3 Tujuan dan Manfaat	6
1.3.1 Tujuan	6
1.3.2 Manfaat	6
1.4 Metode Penelitian	6
1.4.1. Metode Studi Pustaka dan Literatur	7
1.4.2. Metode Pembuatan Model.....	7
1.4.3. Metode Pengujian	7
1.4.4. Metode Analisis dan Kesimpulan	7
1.5 Sistematika Penulisan	8
BAB II TINJAUAN PUSAKA	10
2.1 Penelitian Terkait.....	10
2.2 <i>Cloud Computing</i>	22
2.2.1 Service Model <i>Cloud Computing</i>	23
2.2.2 Jenis Cloud.....	24
2.2.3 Komponen <i>Cloud Computing</i>	25
2.3 <i>Reinforcement Learning</i>	25
2.3.1 Elemen-elemen dari <i>Reinforcement Learning</i>	27

2.3.2 Q-Learning.....	30
2.3.3 Function Approximation dengan Deep Neural Network.....	30
2.3.4 Deep Q-Network (DQN)	31
2.4 <i>Proximal Policy Optimization</i>	32
2.4.1 <i>Policy Gradient Methods</i>	32
2.4.2 Policy Gradient Theorem.....	32
2.4.3 <i>Proximal Policy Optimization (PPO)</i>	33
BAB III METODELOGI PENELITIAN.....	35
3.1 Pendahuluan.....	35
3.2 Kerangka Kerja Penelitian	35
3.3 Dataset	37
3.4 Load Dataset	38
3.5 Preprocessing	40
3.6 Visualisasi Data.....	40
3.7 Inisialisasi Environment Class.....	43
3.8 Training Model PPO dan DQN.....	44
3.9 Evaluasi Model PPO dan DQN	45
3.10 Visualisasi Data.....	46
BAB IV HASIL DAN PEMBAHASAN	48
4.1 Persiapan Sistem.....	48
4.2 Analisis Dataset dengan Pendekatan <i>Reinforcement Learning</i>	49
4.2.1 Load Dataset	49
4.2.2 Pre-Processing	50
4.2.3 Visualisasi Data	53
4.2.4 Deep Q Learning	55
4.2.5 <i>Proximal Policy Optimization</i>	66
4.2.6 Perbandingan <i>Reinforcement Learning</i> DQN dan PPO.....	73
4.2.7 Visualisasi Dataset	78
BAB V KESIMPULAN DAN SARAN	97
5.1. Kesimpulan	97
5.2. Saran	97
DAFTAR PUSTAKA	99
LAMPIRAN	

DAFTAR GAMBAR

Gambar 2.1 Timeline Penelitian Terkait.....	21
Gambar 2.2 Aspek – aspek standar <i>Cloud</i> [30]	22
Gambar 2.3 Komponen <i>Cloud Computing</i> [33]	25
Gambar 2.4 Skenario <i>Reinforcement Learning</i> [35].....	26
Gambar 2.5 Struktur Formula DQN Model.....	31
Gambar 2.6 PPO Model Formula	33
Gambar 3.1 Kerangka Penelitian.....	36
Gambar 3.2 Framework Dataset.....	37
Gambar 3.3 Framework <i>Reinforcement Learning</i>	39
Gambar 3.4 Arsitektur Cloud Computing.....	39
Gambar 3.5 Arsitektur Environment.....	41
Gambar 4.1 Dataset	49
Gambar 4.2 Hasil konversi dataset	51
Gambar 4.3 Hasil Encoding Label.....	52
Gambar 4.4 Hasil <i>Cleaning Data</i>	53
Gambar 4.5 Hasil Visualisasi Data dengan <i>Donut Chart</i>	54
Gambar 4.6 Pemeriksaan data final sebelum training	55
Gambar 4.8 Visualisasi hasil prediksi dalam <i>Plot Batang</i>	60
Gambar 4.9 Visualisasi Hasil Jumlah Status setelah DQN.....	61
Gambar 4.10 Heatmap Q-Table	62
Gambar 4.11 Nilai Q untuk State dengan Indeks 10	63
Gambar 4.12 <i>Confusion Matrix</i> model DQN	63
Gambar 4.13 Visualisasi hasil Evaluasi Model PPO	70
Gambar 4.14 <i>Confusion Matrix</i> model PPO.....	71
Gambar 4.15 Visualisasi Jaringan Hubungan Antar Komponen menggunakan Circos.....	89
Gambar 4.16 Visualisasi Jaringan Hubungan Antar Komponen menggunakan Circos.....	90
Gambar 4.17 Visualisasi Jaringan menggunakan SankeyMatic	92

DAFTAR TABEL

Tabel 2.1 Penelitian Terkait.....	12
Tabel 4.1 Tabel <i>Grid Search Hyper Parameter</i> DQN	73
Tabel 4.2 Tabel <i>Grid Search Hyper Parameter</i> PPO	74
Tabel 4.3 Tabel Perbandingan Hasil Analisis	75
Tabel 4.4 Tabel Perbandingan Evaluasi DQN dan PPO.....	77
Tabel 4.5 Tabel Perbandingan Hasil	78
Tabel 4.6 Tabel Perbandingan Persentase label	83

DAFTAR LAMPIRAN

- Lampiran 1. Script Phyton Deep Q-Learning
- Lampiran 1. Script Phyton Deep Q-Learning
- Lampiran 2 Script Phyton Policy Optimization
- Lampiran 3. Hasil Cek Plagiarisme di Turnitin Halaman Judul
- Lampiran 4. Hasil Presentase Cek Plagiarisme di Turnitin
- Lampiran 5 Lembar Keterangan Pengecekan Similiarity
- Lampiran 6 Form Revisi Penguji Skripsi
- Lampiran 5 Form Revisi Dosen Pembimbing I
- Lampiran 6 Form Revisi Dosen Pembimbing II

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam era digital saat ini, *Cloud Computing* telah menjadi tren yang semakin populer dan diadopsi secara luas oleh berbagai organisasi dan perusahaan. Salah satu layanan utama yang disediakan oleh *cloud computing* adalah *Infrastructure as a Service* (IaaS), yang memungkinkan pengguna untuk mengakses dan mengatur sumber daya komputasi seperti penyimpanan, jaringan, dan sumber daya virtualisasi lainnya secara on-demand [1]. *Cloud computing*, khususnya IaaS, menjadi tren adalah skalabilitas, efisiensi biaya, waktu deployment yang cepat, serta keandalan dan ketersediaan yang tinggi. IaaS memungkinkan organisasi untuk dengan mudah menyesuaikan sumber daya komputasi sesuai kebutuhan, baik dengan menambah atau mengurangi kapasitas. Selain itu, IaaS mengadopsi model pembayaran per penggunaan (*pay-as-you-go*), sehingga organisasi hanya membayar sumber daya yang mereka gunakan. Selain itu IaaS memiliki fleksibilitas dan skalabilitas yang tinggi, waktu deployment yang cepat, serta akses global dan sumber daya komputasi yang hampir tak terbatas.

Peningkatan adopsi IaaS, menurut laporan dari Gartner pada tahun 2022, pada pasar IaaS global tumbuh sebesar 37,4% dibandingkan tahun sebelumnya, mencapai total pendapatan sebesar \$90,9 miliar. Angka ini menunjukkan peningkatan adopsi IaaS yang signifikan di seluruh dunia. Dengan meningkatnya adopsi IaaS, optimalisasi performa layanan menjadi sangat penting untuk memastikan kualitas layanan yang baik dan efisiensi sumber daya. Namun, mengelola sumber daya IaaS secara optimal bukanlah tugas yang mudah, terutama dengan adanya fluktuasi permintaan dan beban kerja yang dinamis [2].

Beberapa penelitian terdahulu yang relevan dengan penelitian ini, pertama oleh Sonia *et. al.* [3] dimana pada penelitian dilakukan perbandingan mendetail dari berbagai platform cloud, mencakup aspek-aspek seperti arsitektur, jaringan, platform yang didukung, hypervisor, dan bahasa pemrograman. Hasil penelitian menunjukkan bahwa kinerja implementasi OpenStack 3-tier jauh lebih baik dibandingkan dengan metode

penyebaran otomatis menggunakan DevStack. Eksperimen yang dilakukan pada kedua alat dengan perangkat keras dan skenario uji yang sama menghasilkan hasil yang berbeda. Penelitian selanjutnya oleh Sajjad *et, al* . [4] hasil dari penelitian ini menunjukkan bahwa pentingnya manajemen sumber daya cloud yang elastis dan isolasi mesin virtual serta jaringan untuk meningkatkan kinerja cloud. Hal ini karena sejumlah mesin menggunakan hard drive dan jaringan yang sama yang dapat berdampak pada kinerja cloud. Ketika banyak pengguna mentransfer data, sumber daya ini terlalu banyak digunakan sehingga mengakibatkan kinerja cloud menurun.

Reinforcement Learning (RL) adalah salah satu cabang machine learning yang berpotensi untuk mengatasi tantangan untuk belajar dari interaksi dengan lingkungan dan menyesuaikan alokasi sumber daya secara dinamis. RL memungkinkan agen (dalam hal ini, sistem manajemen IaaS) untuk belajar dan mengambil tindakan optimal berdasarkan interaksi dengan lingkungan dan umpan balik yang diterima. Dengan menggunakan RL, sistem manajemen IaaS dapat mengoptimalkan alokasi sumber daya secara dinamis dan responsif terhadap perubahan permintaan, sehingga meningkatkan efisiensi dan performa layanan secara keseluruhan[5].

Penelitian terdahulu dalam *cloud computing* masih memiliki kelemahan terkait perbandingan performa platform, manajemen sumber daya yang elastis, serta mekanisme isolasi sumber daya. Penelitian oleh Sonia *et al.* menunjukkan bahwa OpenStack 3-tier memiliki kinerja yang lebih baik dibandingkan DevStack, dengan latensi lebih rendah (150 ms vs. 220 ms), throughput lebih tinggi (2000 TPS vs. 1500 TPS), serta waktu respons yang lebih cepat. Sajjad *et al.* menemukan bahwa sumber daya cloud seperti jaringan dan disk mengalami penurunan kinerja hingga 35% ketika terlalu banyak VM berbagi sumber daya, khususnya pada transfer data.

Metode *Reinforcement Learning* berpotensi mengatasi permasalahan tersebut dengan kemampuannya belajar dan beradaptasi secara dinamis berdasarkan interaksi dengan lingkungan cloud yang kompleks. Dengan pendekatan *Reinforcement Learning* (RL), sistem manajemen IaaS dapat mengurangi penggunaan CPU hingga 15%, menurunkan latensi 35%, dan meningkatkan efisiensi memori 20%, serta memastikan alokasi sumber daya lebih responsif dan stabil. Melalui proses penghargaan dan hukuman, agen *Reinforcement Learning* dapat menemukan strategi optimal dalam

mengalokasikan sumber daya secara efisien, elastis, dan terisolasi. Eksplorasi dan eksploitasi yang menjadi ciri khas *Reinforcement Learning* memungkinkan agen mencoba berbagai kombinasi tindakan dan mempelajari pola data untuk mengambil tindakan tepat. Penggabungan dengan algoritma lain juga dapat membantu menyelesaikan masalah penjadwalan dan alokasi sumber daya yang lebih kompleks.

Penelitian lainnya yang telah menggunakan *Reinforcement Learning* oleh Lei Yu *et, al.* [6] hasil penelitian menunjukkan bahwa penerapan algoritma DQN dengan *Reinforcement Learning* yang ditingkatkan pada masalah penjadwalan layanan gabungan dalam lingkungan *cloud* kontainer dapat secara efektif mengurangi waktu penyelesaian layanan gabungan. Sementara itu, metode ini juga meningkatkan *Quality of Service(QoS)* dan pemanfaatan sumber daya di lingkungan *cloud* kontainer. Penelitian lainnya Umer Ahmed Butt *et, al* [7] oleh penelitian ini meninjau berbagai algoritma *machine learning* yang digunakan untuk mengatasi masalah keamanan *cloud*, hasilnya menunjukkan bahwa *Reinforcement Learning* (RL) tergolong sebagai salah satu dari tiga model utama *machine learning*. Tantangan utama yang muncul dalam *Reinforcement Learning*, yang tidak ditemui dalam jenis *machine learning* lainnya adalah perluasan kemampuan antara eksplorasi dan eksploitasi. Dari berbagai pendekatan ML yang ada, RL adalah yang paling mendekati cara belajar manusia.

Penelitian terdahulu lainnya oleh Guangyou Zhou *et, al.* [8] memperoleh hasil bahwa *Deep Reinforcement Learning* memiliki potensi besar dalam penjadwalan *cloud* dan membuka banyak arah penelitian baru. Salah satu tantangan utama adalah bagaimana menggabungkan *Reinforcement Learning* dengan berbagai jenis algoritma lain untuk menyelesaikan masalah penjadwalan yang lebih kompleks. Hal ini termasuk dalam pemilihan algoritma penjadwalan yang optimal. Penelitian dan eksplorasi lebih lanjut dalam bidang ini masih sangat diperlukan.

Meskipun penelitian terdahulu telah memberikan kontribusi penting dalam penerapan *Reinforcement Learning* untuk penjadwalan layanan *cloud computing*, namun masih terdapat beberapa kekurangan seperti keterbatasan dalam menangani ruang aksi besar dan kontinyu, tantangan keseimbangan eksplorasi-eksploitasi, serta kesulitan menggabungkan dengan algoritma lain untuk masalah kompleks. Oleh karena itu, penelitian ini mengusulkan penggunaan algoritma PPO (*Proximal Policy*

Optimization) yang dirancang untuk mengatasi kekurangan-kekurangan tersebut. PPO memiliki kemampuan menangani ruang aksi besar, menjaga keseimbangan eksplorasi-eksploitasi dengan baik, serta mudah digabungkan dengan teknik deep learning dan optimasi lainnya, sehingga berpotensi memberikan solusi penjadwalan yang lebih optimal dan meningkatkan performa layanan cloud computing (IaaS) secara keseluruhan.

Penelitian ini membedakan dirinya dari penelitian terdahulu dengan menerapkan algoritma *Proximal Policy Optimization* (PPO) dan Deep Q-Network (DQN) dalam konteks Infrastructure as a Service (IaaS) pada cloud computing. Sementara penelitian sebelumnya mungkin telah mengeksplorasi berbagai metode RL, penelitian ini khusus menyoroti penerapan PPO dan DQN untuk mengatasi tantangan dalam penjadwalan dan alokasi sumber daya yang kompleks. PPO dirancang untuk menangani ruang aksi besar dan menjaga keseimbangan eksplorasi-eksploitasi dengan baik, sementara DQN menggabungkan kekuatan deep learning dengan Q-learning tradisional untuk keputusan yang lebih optimal. Penelitian ini juga mengatasi kekurangan metodologi RL sebelumnya, seperti keterbatasan dalam menangani ruang aksi besar dan kesulitan integrasi dengan algoritma lain. Dengan menggunakan dataset terbaru dan metode evaluasi komprehensif, penelitian ini menawarkan kontribusi baru dalam praktik manajemen IaaS dan pengetahuan teori RL.

Selain itu RL yang telah menunjukkan keberhasilan dalam menangani masalah optimasi yang kompleks adalah *Deep Q-Network* (DQN). DQN menggabungkan kekuatan *deep learning* dengan konsep Q-learning tradisional, memungkinkan agen untuk belajar kebijakan optimal dalam lingkungan dengan ruang keadaan yang besar dan kompleks. Dalam konteks IaaS, DQN dapat membantu sistem manajemen untuk membuat keputusan yang lebih baik dalam alokasi sumber daya, penjadwalan tugas, dan optimalisasi performa berdasarkan pengalaman dan pembelajaran yang berkelanjutan. Maka dari itu penulis melakukan penelitian ini untuk mengembangkan metode penjadwalan yang lebih efisien dan adaptif dalam lingkungan IaaS *cloud computing*. Dengan memanfaatkan *Reinforcement Learning* menggunakan algoritma DQN dan PPO, diharapkan dapat dicapai peningkatan kinerja, QoS, dan pemanfaatan sumber daya

yang lebih baik, serta solusi untuk tantangan yang belum terpecahkan dalam penjadwalan dan manajemen sumber daya *cloud*.

Penelitian ini mengusulkan penggunaan *Reinforcement Learning* (RL) untuk mengoptimalkan kinerja layanan *Infrastructure as a Service* (IaaS) pada cloud computing. RL merupakan cabang dari *machine learning* yang memungkinkan agen (dalam hal ini, sistem manajemen IaaS) untuk belajar dan mengambil tindakan optimal berdasarkan interaksi dengan lingkungan dan umpan balik yang diterima. Dengan menggunakan RL, sistem manajemen IaaS dapat mengoptimalkan alokasi sumber daya secara dinamis dan responsif terhadap perubahan permintaan, sehingga meningkatkan efisiensi dan performa layanan secara keseluruhan. Hal ini sangat penting karena mengelola sumber daya IaaS secara optimal bukanlah tugas yang mudah, terutama dengan adanya fluktuasi permintaan dan beban kerja yang dinamis. Oleh karena itu, penelitian ini bermaksud untuk mengeksplorasi penggunaan RL dalam konteks IaaS dan optimalisasi performa secara spesifik, karena walaupun beberapa penelitian terdahulu telah mengeksplorasi penggunaan RL dalam konteks manajemen sumber daya *cloud computing*, masih terdapat ruang untuk eksplorasi lebih lanjut.

Dari uraian diatas maka penulis mengambil judul sebagai tugas akhir, yaitu dengan judul “**Analisa Kinerja Infrastructure As A Service Pada Cloud Computing Dengan Pendekatan *Reinforcement Learning***”.

1.2 Perumusan dan Batasan Masalah

Berdasarkan uraian latar belakang diatas maka rumusan masalah pada penelitian adalah bagaimana kinerja performa layanan *Infrastructure as a Service* (IaaS) pada *cloud computing* dengan menggunakan metode *Reinforcement Learning* (RL).

1.2.1 Batasan Masalah

Untuk membatasi ruang lingkup penelitian dan memfokuskan pada permasalahan utama, beberapa batasan masalah yang akan diterapkan adalah:

- 1 Penelitian ini akan berfokus pada layanan IaaS pada cloud computing.
- 2 Metrik performa yang akan dioptimalkan pada penggunaan sumber daya seperti CPU, memori, bandwidth, dan waktu respons, serta ketersediaan layanan.

- 3 Algoritma RL yang akan digunakan adalah *Deep Q-Network (DQN)* dan *Proximal Policy Optimization (PPO)*.

1.3 Tujuan dan Manfaat

1.3.1 Tujuan

Tujuan dari penelitian ini adalah sebagai berikut:

1. Merancang lingkungan (environment) yang merepresentasikan skenario nyata dari layanan IaaS pada cloud computing.
2. Mengimplementasikan agen *Reinforcement Learning (RL)* yang mampu belajar dari lingkungan dan mengambil tindakan optimal untuk mengoptimalkan performa layanan IaaS.
3. Mengevaluasi kinerja agen RL dalam mengoptimalkan performa layanan IaaS berdasarkan metrik-metrik seperti penggunaan sumber daya, ketersediaan layanan, dan waktu respons.

1.3.2 Manfaat

Manfaat dari penelitian ini antara lain:

1. Menghasilkan sistem yang mampu mengalokasikan dan memanfaatkan sumber daya secara dinamis, mengurangi pemborosan, dan meningkatkan efisiensi operasional.
2. Meningkatkan kemampuan adaptasi real-time terhadap perubahan kondisi lingkungan cloud dan kebutuhan pengguna, memastikan layanan berkinerja tinggi dengan latensi rendah, serta mengembangkan model keputusan yang lebih baik dalam menghadapi situasi kompleks dan dinamis pada pengelolaan sumber daya cloud.
3. Menyediakan strategi penjadwalan dan pengambilan keputusan yang optimal untuk meminimalkan waktu penyelesaian layanan yang kompleks dengan biaya operasional yang rendah.

1.4 Metode Penelitian

Pada tugas akhir ini menggunakan metodologi sebagai berikut :

1.4.1. Metode Studi Pustaka dan Literatur

Pentingnya metode ini adalah untuk menggali informasi yang berkaitan dengan optimalisasi performa layanan IaaS pada cloud computing dan metode *Reinforcement Learning*. Kegiatan yang dilakukan meliputi mencari referensi seperti buku, jurnal, artikel ilmiah, dan sumber-sumber relevan lainnya yang dapat digunakan dalam penelitian ini.

1.4.2. Metode Pembuatan Model

Pada metode ini, dilakukan perancangan model untuk lingkungan (*environment*) yang merepresentasikan skenario nyata dari layanan IaaS pada *cloud computing*. Model ini akan digunakan dalam pembuatan *flowchart* dan implementasi sistem.

1.4.3. Metode Pengujian

Metode ini melakukan pengujian terhadap agen *Reinforcement Learning* yang telah diimplementasikan. Pengujian dilakukan untuk mengevaluasi kinerja agen dalam mengoptimalkan performa layanan IaaS. Pengujian akan menghasilkan nilai akurasi dan metrik-metrik lain yang dapat digunakan untuk menganalisis kinerja agen.

1.4.4. Metode Analisis dan Kesimpulan

Hasil dari pengujian pada penelitian ini akan dianalisis untuk mengetahui kekurangan dan kelebihan dari pendekatan yang digunakan. Analisis akan dilakukan terhadap nilai akurasi, efisiensi penggunaan sumber daya, ketersediaan layanan, dan waktu respons yang dihasilkan oleh agen *Reinforcement Learning*. Berdasarkan analisis tersebut, akan dibuat kesimpulan dan rekomendasi untuk penelitian selanjutnya dalam bidang optimalisasi performa layanan IaaS pada *cloud computing* dengan menggunakan metode *Reinforcement Learning*.

1.5 Sistematika Penulisan

Dalam penyusunan tugas akhir, penulisan akan disusun secara sistematis dengan urutan per-bab. Setiap bab akan berisikan sub-bab yang menjelaskan secara detail topik yang bersangkutan. Sistematika penulisan tersusun sebagai berikut:

BAB I. PENDAHULUAN

Pada bab ini akan menjelaskan sub-bab seperti latar belakang, perumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

BAB II. TINJAUAN PUSTAKA

Bab ini akan menjelaskan sub-bab seperti penelitian terkait atau penelitian terdahulu yang relevan, ringkasan hasil kajian literatur, dan landasan teori yang mendukung penelitian ini, seperti konsep *cloud computing*, *Infrastructure as a Service (IaaS)*, *Reinforcement Learning*, dan topik-topik lain yang terkait.

BAB III. METODE PENELITIAN

Pada bab ini akan dijelaskan sub-bab seperti metode pengumpulan data, lingkungan dan spesifikasi perangkat keras serta perangkat lunak yang digunakan, rancangan blok diagram sistem, metode yang digunakan (seperti *Reinforcement Learning*, *Deep Q-Network*, *Proximal Policy Optimization*, dll.), serta *flowchart* dari sistem yang diimplementasikan.

BAB IV. HASIL DAN PEMBAHASAN

Bab ini akan menjelaskan sub-bab seperti analisis dari penelitian yang dilakukan, hasil pengujian dan evaluasi kinerja agen *Reinforcement Learning* dalam mengoptimalkan performa layanan IaaS, serta pembahasan mengenai kelebihan dan kekurangan dari pendekatan yang digunakan.

BAB V. PENUTUP

Pada bab terakhir ini akan menjelaskan sub-bab seperti kesimpulan dari penelitian yang telah dilakukan, serta saran dan rekomendasi untuk penelitian selanjutnya dalam bidang optimalisasi performa layanan IaaS pada *cloud computing* dengan menggunakan metode *Reinforcement Learning* atau metode lain yang relevan.

DAFTAR PUSTAKA

- [1] R. N. Calheiros, E. Masoumi, R. Ranjan, and R. Buyya, “Reinforcement Learning for Virtualized Cloud Resource Management,” *IEEE Transactions on Cloud Computing*, vol. 7, no. 4, p. 10211034, 2019.
- [2] J. Xu, Z. Lyu, and X. Fan, “Reinforcement Learning for Resource Provisioning in Cloud Computing,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 30, no. 12, pp. 2614–2627, 2019.
- [3] S. Shahzadi, Z. U. Qayyum, and T. Dagiuklas, “Infrastructure as a Service (IaaS): A Comparative Performance Analysis of Open-Source Cloud Platforms,” *IEEE International Workshop on Computer Aided Modeling and Design of Communication Links and Networks*, pp. 1–6, 2017, doi: <https://doi.org/10.1109/CAMAD.2017.8031522>.
- [4] M. Sajjad, A. Ali, and A. S. Khan, “Performance Evaluation of Cloud Computing Resources,” 2018. [Online]. Available: www.ijacsa.thesai.org
- [5] Y. Zhai, X. Cheng, and R. Lu, “Reinforcement Learning for Cloud Resource Allocation: A Comprehensive Review,” *IEEE Transactions on Cloud Computing*, vol. 8, no. 4, pp. 1236–1249, 2020.
- [6] L. Yu, P. S. Yu, Y. Duan, and H. Qiao, “A resource scheduling method for reliable and trusted distributed composite services in cloud environment based on deep Reinforcement Learning,” *Front Genet*, vol. 13, Oct. 2022, doi: 10.3389/fgene.2022.964784.
- [7] U. A. Butt *et al.*, “A review of machine learning algorithms for cloud computing security,” Sep. 01, 2020, *MDPI AG*. doi: 10.3390/electronics9091379.
- [8] G. Zhou, W. Tian, R. Buyya, R. Xue, and L. Song, “Deep Reinforcement Learning-based methods for resource scheduling in cloud computing: a review and future directions,” *Artif Intell Rev*, vol. 57, no. 5, May 2024, doi: 10.1007/s10462-024-10756-9.
- [9] Q. Duan, “Cloud service performance evaluation: status, challenges, and opportunities – a survey from the system modeling perspective,” *ScienceDirect*, vol. 3, no. 2, pp. 101–111, May 2017, doi: 10.1016/j.dcan.2016.12.002.
- [10] L. Yu, P. S. Yu, Y. Duan, and H. Qiao, “A resource scheduling method for reliable and trusted distributed composite services in cloud environment based on deep Reinforcement Learning,” *PubMed*, vol. 13, pp. 1–15, Oct. 2022, doi: 10.3389/fgene.2022.964784.
- [11] U. A. Butt *et al.*, “A review of machine learning algorithms for cloud computing security,” Sep. 01, 2020, *MDPI AG*. doi: 10.3390/electronics9091379.

- [12] G. Zhou, W. Tian, and R. Buyya, “Deep *Reinforcement Learning*-based Methods for Resource Scheduling in Cloud Computing: A Review and Future Directions,” *Semantic*, vol. 1, no. 1, pp. 1–18, May 2021, [Online]. Available: <http://arxiv.org/abs/2105.04086>
- [13] T. Zheng, J. Wan, J. Zhang, and C. Jiang, “Deep *Reinforcement Learning*-Based Workload Scheduling for Edge Computing,” *Journal of Cloud Computing*, vol. 11, no. 1, Dec. 2022, doi: 10.1186/s13677-021-00276-0.
- [14] Y. Jin, M. Bouzid, D. Kostadinov, and A. Aghasaryan, “Model-free resource management of cloud-based applications using *Reinforcement Learning*,” *IEEE*, pp. 1–6, Jun. 2018, doi: 10.1109/ICIN.2018.8401615.
- [15] A. Belgacem, S. Mahmoudi, and M. Kihl, “Intelligent multi-agent *Reinforcement Learning* model for resources allocation in cloud computing,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 2391–2404, Jun. 2022, doi: 10.1016/j.jksuci.2022.03.016.
- [16] S. Swarup, E. M. Shakshuki, and A. Yasar, “Task scheduling in cloud using deep *Reinforcement Learning*,” in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 42–51. doi: 10.1016/j.procs.2021.03.016.
- [17] L. Ismail and H. Materwala, “Machine Learning-based Energy-Aware Offloading in Edge-Cloud Vehicular Networks,” in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 328–336. doi: 10.1016/j.procs.2021.07.044.
- [18] G. Liu, G. Chen, and V. Huang, “Policy ensemble gradient for continuous control problems in deep *Reinforcement Learning*,” *Neurocomputing*, vol. 548, Sep. 2023, doi: 10.1016/j.neucom.2023.126381.
- [19] T. Oudaa, H. Gharsellaoui, and S. Ben Ahmed, “An agent-based model for resource provisioning and task scheduling in cloud computing using DRL,” in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 3795–3804. doi: 10.1016/j.procs.2021.09.154.
- [20] S. Pérez, P. Arroba, and J. M. Moya, “Energy-conscious optimization of Edge Computing through Deep *Reinforcement Learning* and two-phase immersion cooling,” *Future Generation Computer Systems*, vol. 125, pp. 891–907, Dec. 2021, doi: 10.1016/j.future.2021.07.031.
- [21] K. Y. Chan *et al.*, “Deep Neural Networks in the Cloud: Review, Applications, Challenges and Research Directions,” *IEE Neurocomputing*, vol. 1, no. 545, pp. 1–24, 2023.
- [22] A. Zeroual, M. Amroune, M. Derdour, and A. Bentahar, “Lightweight deep learning model to secure authentication in Mobile Cloud Computing,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 9, pp. 6938–6948, Oct. 2022, doi: 10.1016/j.jksuci.2021.09.016.

- [23] S. S. Panwar, M. M. S. Rauthan, V. Barthwal, N. Mehra, and A. Semwal, "Machine learning approaches for efficient energy utilization in cloud data centers," in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 1782–1792. doi: 10.1016/j.procs.2024.04.169.
- [24] T. Thein, M. M. Myo, S. Parvin, and A. Gawanmeh, "Reinforcement Learning based methodology for energy-efficient resource allocation in cloud data centers," *Journal of King Saud University - Computer and Information Sciences*, vol. 32, no. 10, pp. 1127–1139, Dec. 2020, doi: 10.1016/j.jksuci.2018.11.005.
- [25] A. Salim, Juliandry, L. Raymond, and J. V. Moniaga, "General pattern recognition using machine learning in the cloud," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 565–570. doi: 10.1016/j.procs.2022.12.170.
- [26] A. Sharma and U. K. Singh, "Modelling of smart risk assessment approach for cloud computing environment using AI & supervised machine learning algorithms," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 243–250, Jun. 2022, doi: 10.1016/j.gltp.2022.03.030.
- [27] K. M. Hosny, A. I. Awad, W. Said, M. Elmezain, E. R. Mohamed, and M. M. Khashaba, "Enhanced whale optimization algorithm for dependent tasks offloading problem in multi-edge cloud computing," *Alexandria Engineering Journal*, vol. 97, pp. 302–318, Jun. 2024, doi: 10.1016/j.aej.2024.04.038.
- [28] H. Wintolo and S. D. Paradita, "Layanan Cloud Computing Berbasis Infrastruktur As a Service Menggunakan Android," *Compiler*, vol. 4, no. 2, pp. 11–18, 2019.
- [29] S. B. Shaw and I. K. Singh, "A survey on cloud computing," 2020.
- [30] K. A.-M. S. ; M. L. A. Munir, *Handbook of Research on Security Considerations in Cloud Computing A volume in the Advances in Information Security, Privacy, and Ethics (AISPE)*, Book Series. 2015.
- [31] S. Singh, Y. S. Jeong, and J. H. Park, "A survey on cloud computing security: Issues, threats, and solutions," *J. Netw. Comput. Appl.*, vol. 75, no. 1, pp. 200–222, 2021.
- [32] J. Sihotang, *Cloud Computing Fundamental*. 2011.
- [33] A. T. Velte, *Cloud Computing*. New York: A Practical Approach, 2010.
- [34] "Reinforcement Learning Toolbox™ User's Guide R2021a," Dec. 2021.
- [35] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, Second. Cambridge: MIT Press, 2018.
- [36] C. J. Watkins and P. Dayan, *Q-learning. Machine learning*, vol. 8. 2019.

- [37] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, and M. G. Bellemare, “Human-level control through deep *Reinforcement Learning*,” *Nature*, vol. 5, no. 7, pp. 529–533, 2021.
- [38] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “*Reinforcement Learning: A survey*,” *Journal of artificial intelligence research*, vol. 4, no. 1, pp. 237–285, 2019.
- [39] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge: MIT Press, 2018.
- [40] L. Engstrom *et al.*, “Implementation Matters in Deep Policy Gradients: A Case Study on PPO and TRPO,” May 2020, [Online]. Available: <http://arxiv.org/abs/2005.12729>