

## **BAB 1**

### **PENDAHULUAN**

#### **1.1.Latar Belakang**

Pengembangan kecerdasan buatan telah menjadi fokus utama dalam berbagai bidang, termasuk permainan komputer, selama dunia digital yang semakin maju. *Reinforcement Learning* (RL), sebuah metode pembelajaran mesin di mana agen belajar membuat keputusan terbaik dengan berinteraksi dengan lingkungannya secara berkelanjutan, RL adalah salah satu pendekatan yang menonjol dalam pengembangan kecerdasan buatan. Dalam RL, agen melakukan tindakan dan menerima sinyal imbalan untuk mengevaluasi seberapa baik tindakan tersebut mencapai tujuan yang diharapkan. Selama proses ini, agen dapat meningkatkan kinerjanya dari waktu ke waktu dengan menyesuaikan strategi dan kebijakan berdasarkan pengalaman yang agen peroleh.

Penelitian ini berfokus pada penggunaan pembelajaran tambahan untuk sistem dalam mengambil keputusan gerakan pada permainan ular, sebuah permainan klasik yang menawarkan tantangan khusus untuk pengembangan kecerdasan buatan. Dalam permainan ini, agen mengendalikan ular untuk memakan makanan dan menghindari bersentuhan dengan dinding atau tubuhnya sendiri. Penelitian ini bertujuan untuk membuat agen kecerdasan buatan yang dapat mengoptimalkan pergerakan ular dan mencapai skor setinggi mungkin dengan menggunakan pustaka *Python* yang disebut *Pygame* dibuat khusus untuk membuat game. *Pygame* memungkinkan pengembang untuk membuat aplikasi multimedia dan game lengkap dengan fitur-fitur grafis, audio, dan input, menggunakan bahasa Python. *Pygame* sangat serbaguna dan kompatibel dengan hampir semua platform dan sistem operasi untuk membangun permainan dasar dan *PyTorch* untuk mengimplementasikan model *Deep Q-Network* (DQN) dengan arsitektur *Linear\_QNet*.

Pada awal implementasi DQN, agen sering kali melakukan gerakan yang tidak efisien. Contohnya, ular bisa terjebak dalam pola gerakan berulang yang tidak efektif, seperti bergerak di lingkaran tanpa henti. Hal ini disebabkan oleh kesalahan dalam menyesuaikan strategi terhadap lingkungan yang berbeda di setiap pengulangannya. Dalam hal menyesuaikan diri dengan lingkungan yang dinamis dan kompleks

dibutuhkan sebutan cara. Menghindari *overfitting*, ini terjadi ketika model kecerdasan buatan terlalu menyesuaikan dengan data pelatihan sehingga kurang efektif dalam kondisi baru saat arena ular belum terdapat sebuah halangan atau rintangan. Mempercepat konvergensi pembelajaran tanpa mengurangi ketepatan prediksi juga penting. Sehingga, penelitian ini tidak hanya berkonsentrasi pada pembuatan algoritma tetapi juga mencakup penerapan metode dalam meningkatkan generalisasi pada gerakan ular agar efektif saat pembelajaran dilakukan oleh agen.

## **1.2.Rumusan Masalah**

Berdasarkan latar belakang masalah dan identifikasi masalah di atas, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana cara membuat basis *snake game* dua dimensi dengan kordinat x,y menggunakan *libraries* pygames.
2. Bagaimana cara agar ular mempelajari cara bergerak, mengenali bahaya, dan memakan makanan.
3. Bagaimana ular mengenali lingkungan sekitarnya dengan menambahkan satu penghalang secara berkala sesuai dengan jumlah makanan yang telah dimakan.

## **1.3.Batasan Masalah**

Batasan ruang lingkup diperlukan agar menjadi parameter untuk menentukan cakupan suatu masalah dalam penelitian ini agar tindakan yang dilakukan tidak menyimpang dari maksud yang sebenarnya. Selain itu membantu mengarahkan fokus pada hal-hal yang relevan dan memperoleh pemahaman yang lebih dalam tentang masalah tersebut. Berdasarkan latar belakang dan Batasan masalah adalah sebagai berikut:

1. Pembuatan sistem lokal kecerdasan buatan agar dapat mengajarkan komputer cara bermain *snake game*.
2. Sistem ini akan melakukan training data dengan melakukan iterasi sesuai kebutuhan.
3. Penerapan sistem akan dilakukan pada basis cpu, untuk meminimalisir beban kerja komputer.
4. Kecerdasan buatan ini akan mampu mengenali dan memperbaiki kesalahan pergerakan ular yang membuat state gamenya menjadi game over.
5. Pembuatan kecerdasan buatan ini akan menggunakan Bahasa pemrograman *python* dan bantuan *library pytorch*.

## 1.4.Tujuan

Tujuan dari Tugas Akhir ini adalah mengimplementasi *reinforcement learning DQN*.

Dengan beberapa tujuan yang diharapkan adalah sebagai berikut:

1. Menjelaskan konsep dasar dan aplikasinya untuk *snake game* dari *Reinforcement Learning*.
2. Menjelaskan konsep yang terjadi dan aplikasinya untuk *snake game* dengan model DQN.
3. Membuat kecerdasan buatan yang mengajarkan cara bermain *snake game*.
4. Menunjukkan runtutan algoritma yang terjadi dari proses pembuatan model DQN, *snake game*, dan *model agent*.
5. Menunjukkan proses *training* data dengan nilai rata rata dari jumlah iterasi dibagi dengan jumlah skor.

## 1.5.State of The Art

Tabel 1.1 State of The Art

Judul Jurnal	Pembahasan
<b>Human-level control through deep reinforcement learning</b> <b>Peneliti</b> Volodymyr Mnih, Koray Kavukcuoglu, David Silver, et al. <b>Tahun</b> 2015 <b>Nama Jurnal</b> Nature	<p><b>Hasil Penitian:</b>            Jurnal ini menekankan kemampuan DQN untuk mengambil taktik rumit dan melakukan modifikasi untuk berbagai permainan tanpa memerlukan bantuan manusia atau modifikasi tambahan. Studi ini juga menyoroti kesulitan dalam perencanaan jangka panjang untuk beberapa permainan, seperti Montezuma's Revenge, yang memerlukan perencanaan yang lebih menyeluruh.</p> <p><b>Alasan Menjadi Tinjauan Penilitian:</b>            merupakan langkah penting menuju pengembangan kecerdasan buatan umum karena menunjukkan kemampuan satu algoritme pembelajaran untuk menjadi mahir dalam berbagai tugas sulit.</p>

<p><b>Proximal Policy Optimization Algorithms</b></p> <p><b>Peneliti</b> John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, Oleg Klimov</p> <p><b>Tahun</b> 2017</p> <p><b>Nama Jurnal</b> arXiv preprint arXiv:1707.06347</p>	<p><b>Hasil Penelitian:</b> Makalah ini menjelaskan bagaimana gradien diperkirakan menggunakan proses stokastik dalam metode gradien kebijakan tradisional, yang memperbarui kebijakan. Namun, metode ini sering kali menyebabkan ketidakstabilan dan efisiensi data yang buruk. Dalam tugas kontrol berkelanjutan (misalnya, penggerak robot), algoritme PPO mengungguli metode yang ada seperti A2C dan ACER dalam hal efisiensi sampel dan waktu konvergensi pada tugas tolak ukur. Selain itu, PPO telah menunjukkan kemanjuran dalam domain diskrit seperti permainan Atari.</p> <p><b>Alasan Menjadi Tinjauan Penilitian:</b> salah satu algoritma pembelajaran penguatan yang dirancang untuk meningkatkan efisiensi dan kesederhanaan metode gradien kebijakan</p>
<p><b>Mastering the game of Go with deep neural networks and tree search</b></p> <p><b>Peniliti:</b> David Silver, Aja Huang, Chris J. Maddison, et al.</p> <p><b>Tahun</b> 2016</p> <p><b>Nama Jurnal</b> Nature</p>	<p><b>Hasil Pengujian:</b> membahas pengembangan AlphaGo, sistem AI yang mencapai kinerja super dalam permainan papan Go yang rumit. Penulis mengembangkan program yang dapat menilai posisi papan secara efisien dan memilih langkah terbaik dengan menggabungkan pembelajaran mendalam dan Pencarian Pohon Monte Carlo (MCTS).</p> <p><b>Alasan Menjadi Tinjauan Penilitian:</b> AlphaGo adalah sumber daya utama untuk penelitian tentang strategi AI dan pengambilan keputusan yang optimal karena teknik pembelajaran penguatan yang digunakan di dalamnya, yang melampaui Go dan mewakili metode yang dapat diterapkan pada skenario pengambilan keputusan lainnya.</p>

<p><b>Reinforcement Learning in Robotics: A Survey</b></p> <p><b>Peniliti:</b> Jens Kober, J. Andrew Bagnell, Jan Peters</p> <p><b>Tahun</b> 2013</p> <p><b>Nama Jurnal</b> International Journal of Robotics Research</p>	<p><b>Hasil Pengujian:</b> Jurnal ini merupakan survei komprehensif tentang penerapan pembelajaran penguatan (RL) dalam robotika. Jurnal ini menekankan potensi dan kesulitan penerapan pembelajaran penguatan (RL) pada desain perilaku robotika yang rumit, dengan membandingkan antara fisika dan matematika. Isu-isu penting yang dibahas dalam survei ini meliputi spesifikasi hadiah, Untuk membuat RL mudah dipahami dalam robotika, makalah ini menyoroti pentingnya representasi, perkiraan, dan pengetahuan sebelumnya yang efisien.</p> <p><b>Alasan Menjadi Tinjauan Penitian:</b> mencakup berbagai taktik, termasuk perkiraan fungsi, kebijakan praterstruktur, dan diskritisasi tindakan status cerdas. Jurnal ini juga menawarkan ringkasan menyeluruh tentang topik penelitian aktif dan aplikasi yang berhasil di bidang ini.</p>
<p><b>Exploration Strategies in Deep Reinforcement Learning: A Survey</b></p> <p><b>Peniliti:</b> Marc G. Bellemare, Yavar Naddaf, Joel Veness, Michael Bowling</p> <p><b>Tahun</b> 2021</p> <p><b>Nama Jurnal</b> IEEE Transactions on Neural Networks and Learning Systems</p>	<p><b>Hasil Pengujian:</b> analisis menyeluruh yang mencakup berbagai pendekatan terhadap eksplorasi dalam pembelajaran penguatan mendalam (DRL). Analisis ini berfokus pada bagaimana berbagai metode menyeimbangkan eksplorasi (mencoba tindakan baru untuk menemukan hasil yang berharga) dan eksploitasi (memanfaatkan tindakan yang diketahui untuk memaksimalkan manfaat).</p> <p><b>Alasan Menjadi Tinjauan Penitian:</b> mencakup berbagai taktik, termasuk perkiraan fungsi, kebijakan praterstruktur, dan diskritisasi tindakan status cerdas. Buku ini juga menawarkan ringkasan menyeluruh tentang topik penelitian aktif dan aplikasi yang berhasil di bidang ini.</p>

<p><b>Deep Reinforcement Learning: An Overview</b></p> <p><b>Peneliti:</b> Yuxi Li</p> <p><b>Tahun</b> 2022</p> <p>Nama Jurnal Journal of Machine Learning Research (JMLR)</p>	<p><b>Hasil Pengujian:</b> Cabang pembelajaran mesin yang disebut deep reinforcement learning (deep RL) memadukan jaringan saraf dalam dan pembelajaran penguatan. Banyak aplikasi, seperti visi komputer, robotika, permainan, dan pemrosesan bahasa alami, telah meraih kesuksesan besar dengan pembelajaran mendalam.\</p> <p><b>Alasan Menjadi Tinjauan Penilitian:</b> meneliti proses kritis yang penting untuk terciptanya sistem pembelajaran penguatan yang canggih, termasuk memori, perhatian, pembelajaran transfer, dan pembelajaran tanpa pengawasan.</p>
--	---

## **1.6.Sistematika Penulisan**

Sistematika penulisan ini bertujuan untuk mengorganisir dan menyusun informasi dalam sebuah tulisan dan memudahkan untuk memahami isi tulisan dengan baik dan logis. Sistem penulisan laporan penelitian terdiri dari:

### **BAB I**

#### **: PENDAHULUAN**

Bab ini berisi tentang pengenalan terhadap topik yang akan dibahas seperti latar belakang, batasan masalah, tujuan penulisan berupa metode perancangan yang digunakan dan sistematika penulisan

### **BAB II**

#### **: LANDASAN TEORI**

Bab ini membahas tentang teori-teori atau konsep penelitian terkait yang relevan dengan topik yang dibahas dalam tulisan tersebut.

### **BAB III**

#### **: METODOLOGI PENELITIAN**

Bab ini berfokus pada penjelasan tentang metode dan prosedur yang digunakan dalam melakukan penelitian, termasuk langkah-langkah dalam mengumpulkan data, menganalisa data dan mendapatkan hasil yang relevan.

### **BAB IV**

#### **: HASIL DAN PEMBAHASAN**

Bab ini berisi tentang pengujian program yang digunakan berdasarkan metodenya, penjelasan dengan gambar dan analisis pengujian program berdasarkan hasil percobaan yang telah dilakukan.

**BAB V****: PENUTUP**

Bab ini berisi tentang rangkuman keseluruhan tulisan, kesimpulan yang diperoleh serta saran dan rekomendasi untuk penelitian selanjutnya.