

## 1 مقدمه (از تمپلیت اضافه شود)

- انگیزه پژوهش (مراحل توسعه یادگیری تقویتی و ارزش آن در پردازنده‌های کنونی) ؟؟؟ آموزش بازیکن با استفاده از سخت افزارهای مدرن GPU <----- صرفاً توضیحات کوتاه از cuda برتری سرعت پاسخ این روش
- تعریف مسئله
- اهداف و نوآوری
- محتوای گزارش

## 2 پیشینه پژوهش

- ماموریت‌های بین مداری (با تمرکز بر مدارهای لیاپانوف) در این بخش، به طور مختصر به ماموریت‌های بین مداری اشاره می‌شود. در این مورد، ماموریت‌های بین مداری شامل ماموریت‌هایی هستند که فضاپیما را از یک مدار به مدار دیگر منتقل می‌کنند.
- بازی دیفرانسیلی --> پژوهش‌های انجام شده هدایت بر مبنای بازی دیفرانسیلی در این بخش، به طور مختصر به بازی دیفرانسیلی اشاره می‌شود. در این مورد، پژوهش‌های انجام شده هدایت بر مبنای بازی دیفرانسیلی
- یادگیری تقویتی --> در زمینه هدایت در این بخش، به طور مختصر به یادگیری تقویتی اشاره می‌شود. در این مورد، یادگیری تقویتی یک تکنیک یادگیری ماشینی است که به یک عامل اجازه می‌دهد تا از طریق تعامل با محیط، رفتار خود را بهینه کند. پژوهش‌های انجام شده هدایت بر یادگیری تقویتی
- ارتباط بین یادگیری تقویتی و بازی دیفرانسیلی در این بخش، به طور مختصر به ارتباط بین یادگیری تقویتی و بازی دیفرانسیلی اشاره می‌شود. در این مورد، می‌توان گفت که یادگیری تقویتی می‌تواند برای حل مسائل بازی دیفرانسیلی استفاده شود.

## 3 یادگیری تقویتی (یادگیری تقویتی به صورت کامل)

- تاریخچه/ مقدمه‌ای بر یادگیری تقویتی (مراحل توسعه و ارزش آن در پردازنده‌های کنونی) ؟؟؟ در این بخش، تاریخچه یادگیری تقویتی و مفاهیم پایه‌ای آن به طور مختصر بیان می‌شود.
- انواع روش‌های یادگیری تقویتی (اسم فصل باشد؟) ---> روش‌های پیاده‌سازی توضیح داده می‌شود مانند PPO و Q-learning

در این بخش، انواع روش‌های یادگیری تقویتی به طور مختصر بیان می‌شود. در این مورد، روش‌های یادگیری تقویتی عبارتند از:

- Q-learning
- SARSA
- Policy Gradient
- Deep Reinforcement Learning

○ شبکه عصبی

در این بخش، به طور مختصر به شبکه‌های عصبی اشاره می‌شود. در این مورد، شبکه‌های عصبی یک مدل ریاضی برای توصیف عملکرد مغز انسان هستند.

- عامل TD
- عامل PPO
- در این بخش، به طور مختصر به عامل PPO اشاره می‌شود. در این مورد، عامل PPO یک روش یادگیری تقویتی است که از PPO (Proximal Policy Optimization) استفاده می‌کند.
- ...
- 

#### 4 بازی دیفرانسیلی

- مقدمه‌ای بر بازی دیفرانسیلی (یک بخش به جای باشید که ترنزینت باشد؟)
- یادگیری مبتنی بر بازی دیفرانسیلی
- در این بخش، به طور مختصر به یادگیری مبتنی بر بازی دیفرانسیلی اشاره می‌شود. در این مورد، یادگیری مبتنی بر بازی دیفرانسیلی یک تکنیک یادگیری ماشینی است که از بازی دیفرانسیلی برای حل مسائل استفاده می‌کند.

#### 5 طراحی الگوریتم یادگیری تقویتی مبتنی بر بازی دیفرانسیلی

- یادگیری تقویتی مبتنی بر بازی دیفرانسیلی
- طراحی عامل مبتنی بر بازی دیفرانسیلی
  - عامل TD مبتنی بر بازی دیفرانسیلی
  - عامل PPO مبتنی بر بازی دیفرانسیلی

روابط توسعه داده شده به صورت کامل برای روش‌های مختلف مانند TD Q-learning sarsa

## 6 مدل سازی محیط یادگیری سه جسمی

- مکانیزم حرکتی فضاپیما
- دستگاه مختصات
- معادلات حرکتی
- محاسبه مسیر مدار لیاپاتوف
- محاسبه مسیر حرکت بین دو مدار لیاپاتوف
- متغیرهای حالت
- فرمان هدایت
- پاداش مبتنی بر بازی دیفرانسیلی
- جست و جوی نقطه هدف <----<-- الگوریتم بهینه سازی جهت پیدا کردن نقطه هدف

## 7 شبیه سازی عامل در محیط سه جسمی

- آموزش عامل
- زمان و نوع پردازنده برای انواع الگوریتم ها بیان شود.
- بررسی عملکرد عامل در شبیه سازی
- مقایسه نتایج و بررسی نتایج برای حالت های مختلف

## 8 پیاده سازی سخت افزار در حلقه عملکرد عامل = در محیط سه جسمی

## 9 نتیجه گیری

- مقایسه بین روش های ارائه شده
- پیشنهاد برای پروژه های آتی

پیشینه پژوهش با مقدمه ای بر آن فرق دارد؟  
آیا مقدمه بر یادگیری تقویتی و بازی دیفرانسیلی حذف شود؟  
تاریخچه ها کجا باشد؟



1. فصل اول: مقدمه (Introduction)

1.1. انگیزه (Motivation)

1.2. تعریف مسئله (Problem Definition)

1.3. اهداف و نوآوری‌ها (Objectives and Innovation)

1.4. مروری کلی بر یادگیری تقویتی چندعاملی (Brief Overview of MARL)

1.5. ساختار پایان‌نامه (Thesis Structure)

2. فصل دوم: پیشینه پژوهش (Literature Review)

2.1. مرور مطالعات یادگیری تقویتی

2.2. مرور مطالعات یادگیری تقویتی چندعاملی

2.3. مطالعات مرتبط با امنیت در یادگیری تقویتی چندعاملی

2.4. پژوهش‌های مرتبط با شبیه‌سازی عامل‌ها

2.5. مطالعات یادگیری در حلقه انسان

2.6. خلاصه و شکاف‌های موجود در پژوهش‌ها

3. فصل سوم: یادگیری تقویتی (RL)

3.1. مفاهیم پایه (Basic Concepts)

3.2. فرآیند تصمیم‌گیری مارکوف (Markov Decision Process)

3.3. الگوریتم DDPG

3.4. الگوریتم PPO

3.5. الگوریتم SAC

3.6. الگوریتم TD3

3.7. مقایسه الگوریتم‌ها

4. فصل چهارم: یادگیری تقویتی چندعاملی (MARL)

4.1. تعاریف و مفاهیم اساسی (Definitions)

4.2. اهمیت یادگیری تقویتی چندعاملی (Importance)

4.3. نظریه بازی‌ها (Game Theory)

4.4. تعادل نش (Nash Equilibrium)

4.5. امنیت در (MARL Safety in MARL)

4.6. الگوریتم‌های (MARL Algorithms)

4.7. چالش‌های موجود در MARL

5. فصل پنجم: شبیه‌سازی عامل‌ها (Agent Simulation)

5.1. معرفی محیط شبیه‌سازی

5.2. پارامترهای شبیه‌سازی

5.3. مدل‌سازی عامل‌ها

5.4. روش‌های ارزیابی

5.5. نتایج اولیه شبیه‌سازی

6. فصل ششم: یادگیری تقویتی چندعاملی امن (Safe MARL)

6.1. چارچوب پیشنهادی

6.2. معماری سیستم

6.3. الگوریتم پیشنهادی

6.4. مکانیزم‌های امنیتی

6.5. پیاده‌سازی و جزئیات فنی

6.6. ارزیابی و مقایسه

7. فصل هفتم: شبیه‌سازی در حلقه انسان (Human-in-the-Loop)

7.1. معماری سیستم تعاملی انسان-عامل

7.2. روش‌های دریافت بازخورد انسانی

7.3. ادغام بازخورد انسانی در یادگیری عامل‌ها

7.4. آزمایش‌ها و نتایج

7.5. مزایا و چالش‌ها

8. فصل هشتم: نتایج و تحلیل (Results and Analysis)

8.1. معیارهای ارزیابی

8.2. تحلیل نتایج مقایسه‌ای

8.3. تحلیل امنیت و کارایی

8.4. تحلیل تأثیر مداخله انسانی

8.5. بحث و تفسیر نتایج

9. فصل نهم: نتیجه‌گیری و کارهای آینده (Conclusion and Future Work)

9.1. خلاصه دستاوردها

9.2. پاسخ به سؤالات پژوهش

9.3. محدودیت‌های پژوهش

9.4. پیشنهادات برای مطالعات آینده

پیوست (Appendix)

A. جزئیات پیاده‌سازی

B. کدهای برنامه

C. داده‌های اضافی

D. مستندات فنی



## 1. فصل اول: مقدمه (Introduction)

### 1.1. انگیزه (Motivation)

توضیح دهید چرا یادگیری تقویتی چندعاملی امن اهمیت دارد و چه مسائل مهمی در دنیای واقعی را حل می‌کند.

### 1.2. تعریف مسئله (Problem Definition)

مسئله اصلی پژوهش خود را با ذکر چالش‌های موجود در امنیت سیستم‌های چندعاملی به طور دقیق تعریف کنید.

### 1.3. اهداف و نوآوری‌ها (Objectives and Innovation)

اهداف مشخص پژوهش و نوآوری‌های پیشنهادی خود را با تمرکز بر شکاف‌های موجود در پژوهش‌های قبلی بیان کنید.

### 1.4. مروری کلی بر یادگیری تقویتی چندعاملی (Brief Overview of MARL)

مفهوم کلی یادگیری تقویتی چندعاملی را معرفی کرده و تفاوت اصلی آن با یادگیری تقویتی تک‌عاملی را بیان کنید.

### 1.5. ساختار پایان‌نامه (Thesis Structure)

خلاصه‌ای از محتوای هر فصل را ارائه دهید تا خواننده دیدی کلی از جریان منطقی پایان‌نامه داشته باشد.

## 2. فصل دوم: پیشینه پژوهش (Literature Review)

### 2.1. مرور مطالعات یادگیری تقویتی

مهم‌ترین پیشرفت‌ها و روش‌های یادگیری تقویتی مرتبط با موضوع پژوهش را با تمرکز بر کاربردهای مشابه بررسی کنید.

### 2.2. مرور مطالعات یادگیری تقویتی چندعاملی

تحقیقات کلیدی در زمینه MARL را با دسته‌بندی روش‌ها به رویکردهای همکارانه و رقابتی مرور کنید.

### 2.3. مطالعات مرتبط با امنیت در یادگیری تقویتی چندعاملی

روش‌های موجود برای بهبود امنیت در سیستم‌های MARL را با تمرکز بر نقاط قوت و ضعف هر روش بررسی کنید.

### 2.4. پژوهش‌های مرتبط با شبیه‌سازی عامل‌ها

محیط‌های شبیه‌سازی مختلف برای آزمایش الگوریتم‌های MARL را با تأکید بر قابلیت‌ها و محدودیت‌های هر محیط مقایسه کنید.

### 2.5. مطالعات یادگیری در حلقه انسان

روش‌های موجود برای ادغام بازخورد انسانی در سیستم‌های یادگیری تقویتی را با ذکر مزایا و چالش‌های هر روش بررسی کنید.

## 2.6. خلاصه و شکاف‌های موجود در پژوهش‌ها

شکاف‌های موجود در پژوهش‌های قبلی را شناسایی کرده و توضیح دهید چگونه پژوهش شما به پر کردن این شکاف‌ها کمک می‌کند.

## 3. فصل سوم: یادگیری تقویتی (RL)

### 3.1. مفاهیم پایه (Basic Concepts)

مفاهیم اساسی یادگیری تقویتی مانند عامل، محیط، حالت، عمل، پاداش و تابع ارزش را با زبانی ساده و روشن توضیح دهید.

### 3.2. فرآیند تصمیم‌گیری مارکوف (Markov Decision Process)

فرآیند تصمیم‌گیری مارکوف را به عنوان چارچوب ریاضی یادگیری تقویتی با تمرکز بر ویژگی‌های مارکوفی بودن معرفی کنید.

### 3.3. الگوریتم DDPG

الگوریتم DDPG را با تمرکز بر معماری آن، مزایا و محدودیت‌ها برای مسائل با فضای عمل پیوسته توضیح دهید.

### 3.4. الگوریتم PPO

الگوریتم PPO را با تأکید بر مفهوم بهینه‌سازی سیاست نامستقیم و مزایای آن در پایداری یادگیری شرح دهید.

### 3.5. الگوریتم SAC

الگوریتم SAC را با تمرکز بر مفهوم بیشینه‌سازی آنتروپی و نقش آن در اکتشاف کارآمد توضیح دهید.

### 3.6. الگوریتم TD3

الگوریتم TD3 را با تأکید بر بهبودهای آن نسبت به DDPG برای کاهش تخمین بیش از حد شرح دهید.

### 3.7. مقایسه الگوریتم‌ها

الگوریتم‌های معرفی شده را از نظر کارایی، پایداری، پیچیدگی محاسباتی و مناسب بودن برای سناریوهای مختلف مقایسه کنید.

## 4. فصل چهارم: یادگیری چندعاملی (MARL)

### 4.1. تعاریف و مفاهیم اساسی (Definitions)

تعاریف پایه‌ای سیستم‌های چندعاملی و چالش‌های خاص آن‌ها در مقایسه با سیستم‌های تک‌عاملی را توضیح دهید.

### 4.2. اهمیت یادگیری تقویتی چندعاملی (Importance)

اهمیت MARL را با تمرکز بر کاربردهای واقعی آن در زمینه‌هایی مانند خودروهای خودران، ربات‌ها و شبکه‌های هوشمند شرح دهید.

### 4.3. نظریه بازی‌ها (Game Theory)

مفاهیم اساسی نظریه بازی‌ها مانند انواع بازی‌ها، استراتژی‌ها و مفهوم تعادل را در زمینه MARL توضیح دهید.

### 4.4. تعادل نش (Nash Equilibrium)

مفهوم تعادل نش را به عنوان راه‌حلی برای بازی‌های استراتژیک و کاربرد آن در MARL با مثال‌های ساده شرح دهید.

### 4.5. امنیت در (MARL Safety in MARL)

چالش‌های امنیتی خاص سیستم‌های MARL و روش‌های موجود برای مقابله با آن‌ها را بررسی کنید.

### 4.6. الگوریتم‌های (MARL Algorithms)

الگوریتم‌های مهم MARL را با تقسیم‌بندی آن‌ها به دسته‌های همکارانه، رقابتی و ترکیبی معرفی کنید.

### 4.7. چالش‌های موجود در MARL

چالش‌های اصلی MARL مانند انفجار فضای حالت-عمل، هماهنگی بین عامل‌ها و عدم ایستایی محیط را تشریح کنید.

## 5. فصل پنجم: شبیه‌سازی عامل‌ها (Agent Simulation)

### 5.1. معرفی محیط شبیه‌سازی

محیط شبیه‌سازی مورد استفاده در پژوهش را با ذکر ویژگی‌های کلیدی آن معرفی کنید.

### 5.2. پارامترهای شبیه‌سازی

پارامترهای اصلی مورد استفاده در شبیه‌سازی مانند تعداد عامل‌ها، فضای حالت و عمل، و تابع پاداش را با جزئیات شرح دهید.

### 5.3. مدل‌سازی عامل‌ها

معماری و ساختار عامل‌های مورد استفاده در شبیه‌سازی را با ذکر قابلیت‌ها و محدودیت‌های آن‌ها توصیف کنید.

### 5.4. روش‌های ارزیابی

معیارها و روش‌های ارزیابی عملکرد عامل‌ها را برای سنجش میزان موفقیت در دستیابی به اهداف شرح دهید.

### 5.5. نتایج اولیه شبیه‌سازی

نتایج اولیه شبیه‌سازی را برای نشان دادن عملکرد پایه الگوریتم‌ها قبل از اعمال بهبودهای پیشنهادی ارائه دهید.

## 6. فصل ششم: یادگیری تقویتی چندعاملی امن (Safe MARL)

### 6.1. چارچوب پیشنهادی

چارچوب کلی روش پیشنهادی خود برای بهبود امنیت در سیستم‌های MARL را با تمرکز بر نوآوری‌ها معرفی کنید.

## 6.2. معماری سیستم

معماری سیستم پیشنهادی را با ذکر اجزای اصلی و نحوه تعامل آن‌ها در قالب دیاگرام و توضیحات شرح دهید.

## 6.3. الگوریتم پیشنهادی

الگوریتم پیشنهادی خود را با جزئیات کافی برای پیاده‌سازی، همراه با شبه‌کد و توضیحات مرحله به مرحله ارائه دهید.

## 6.4. مکانیزم‌های امنیتی

مکانیزم‌های امنیتی به کار رفته در روش پیشنهادی را با تمرکز بر نحوه مقابله با چالش‌های امنیتی شناسایی شده توضیح دهید.

## 6.5. پیاده‌سازی و جزئیات فنی

جزئیات فنی پیاده‌سازی روش پیشنهادی شامل پارامترها، ساختارهای داده و فناوری‌های مورد استفاده را شرح دهید.

## 6.6. ارزیابی و مقایسه

نتایج ارزیابی روش پیشنهادی را با روش‌های موجود برای نشان دادن میزان بهبود امنیت و کارایی مقایسه کنید.

## 7. فصل هفتم: شبیه‌سازی در حلقه انسان (Human-in-the-Loop)

### 7.1. معماری سیستم تعاملی انسان-عامل

معماری سیستم تعاملی بین انسان و عامل‌ها را با تمرکز بر نحوه دریافت و استفاده از بازخورد انسانی توضیح دهید.

### 7.2. روش‌های دریافت بازخورد انسانی

روش‌های مختلف برای دریافت بازخورد انسانی مانند نمایش ترجیحات، آموزش مستقیم و نظارت را شرح دهید.

### 7.3. ادغام بازخورد انسانی در یادگیری عامل‌ها

نحوه ادغام بازخورد انسانی در فرآیند یادگیری عامل‌ها برای بهبود امنیت و کارایی را با جزئیات توضیح دهید.

### 7.4. آزمایش‌ها و نتایج

آزمایش‌های انجام شده با مداخله انسانی و نتایج آن‌ها را با تمرکز بر میزان بهبود نسبت به حالت بدون مداخله ارائه دهید.

### 7.5. مزایا و چالش‌ها

مزایا و چالش‌های استفاده از بازخورد انسانی در سیستم‌های MARL را با تمرکز بر تعادل بین کارایی و قابلیت استفاده بررسی کنید.

## 8. فصل هشتم: نتایج و تحلیل (Results and Analysis)

### 8.1. معیارهای ارزیابی

معیارهای مورد استفاده برای ارزیابی عملکرد روش پیشنهادی را با توجیه انتخاب هر معیار شرح دهید.

### 8.2. تحلیل نتایج مقایسه‌ای

نتایج آزمایش‌های مقایسه‌ای بین روش پیشنهادی و روش‌های موجود را با استفاده از جداول و نمودارها تحلیل کنید.

### 8.3. تحلیل امنیت و کارایی

نتایج آزمایش‌ها از منظر بهبود امنیت و حفظ کارایی را با تمرکز بر تعادل بین این دو هدف بررسی کنید.

### 8.4. تحلیل تأثیر مداخله انسانی

تأثیر مداخله انسانی بر عملکرد سیستم را با مقایسه نتایج در حالت‌های با و بدون مداخله انسانی تحلیل کنید.

### 8.5. بحث و تفسیر نتایج

نتایج به دست آمده را با دیدگاهی انتقادی تفسیر کرده و معنی عملی آن‌ها را برای کاربردهای واقعی بیان کنید.

## 9. فصل نهم: نتیجه‌گیری و کارهای آینده (Conclusion and Future Work)

### 9.1. خلاصه دستاوردها

خلاصه‌ای از دستاوردهای اصلی پژوهش را با تمرکز بر نوآوری‌ها و میزان دستیابی به اهداف ارائه دهید.

### 9.2. پاسخ به سؤالات پژوهش

به سؤالات اصلی پژوهش که در مقدمه مطرح شده‌اند، با استناد به نتایج به دست آمده پاسخ دهید.

### 9.3. محدودیت‌های پژوهش

محدودیت‌های موجود در پژوهش را با صداقت بیان کرده و تأثیر احتمالی آن‌ها بر نتایج را ارزیابی کنید.

### 9.4. پیشنهادات برای مطالعات آینده

مسیرهای پژوهشی آینده را بر اساس یافته‌ها و شکاف‌های شناسایی شده در طول پژوهش پیشنهاد دهید.

## پیوست (Appendix)

### A. جزئیات پیاده‌سازی

جزئیات فنی پیاده‌سازی که برای درک کامل روش پیشنهادی ضروری است اما در متن اصلی باعث انحراف از موضوع می‌شود را ارائه دهید.

## B. کدهای برنامه

نمونه‌هایی از کدهای اصلی مورد استفاده در پیاده‌سازی را با توضیحات کافی برای درک عملکرد آن‌ها ارائه دهید.

## C. داده‌های اضافی

داده‌های تکمیلی مانند نتایج آزمایش‌های بیشتر، جداول مقایسه‌ای و نمودارهای اضافی را برای پشتیبانی از یافته‌های اصلی ارائه دهید.

## D. مستندات فنی

مستندات فنی مانند راهنمای استفاده از کد، توضیح پارامترها و نحوه اجرای آزمایش‌ها را برای تکرارپذیری پژوهش فراهم کنید.