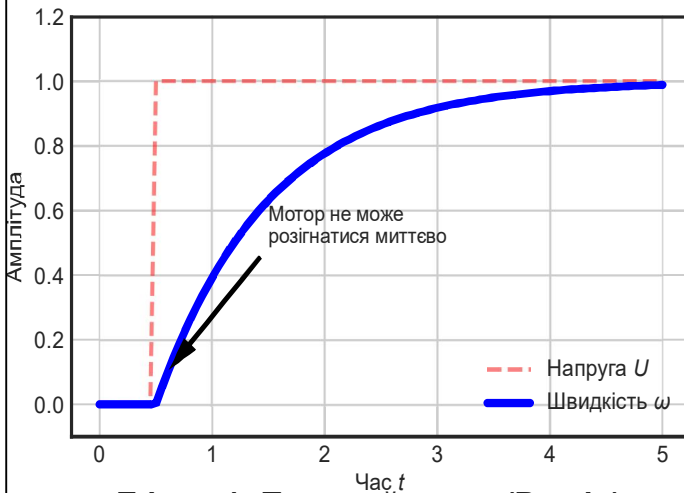
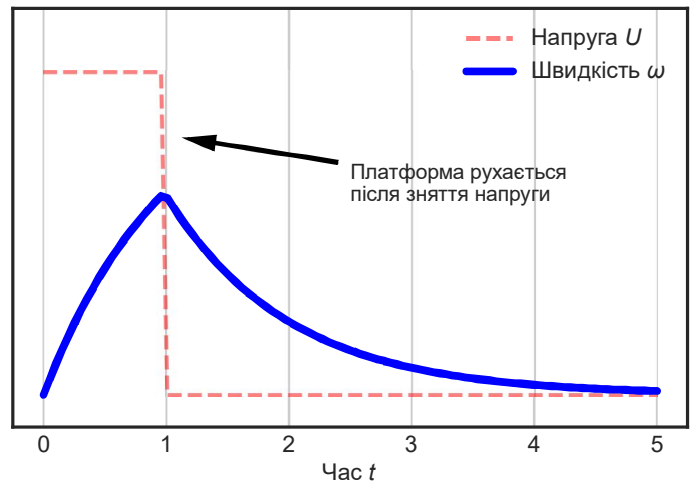


Об'єкт керування: Аперіодична ланка 2-го порядку



Ефект 1: Плавний старт (Розгін)



Ефект 2: Вибіг (Інерційна зупинка)

Фізична природа 2 – го порядку :

Система складається з двох інерційних компонентів:

1. **Двигун** : Струм не зростає миттєво -> Плавний старт.
2. **Платформа (J)** : Масивне тіло не зупиняється миттєво -> Вибіг.

Аперіодичний характер :

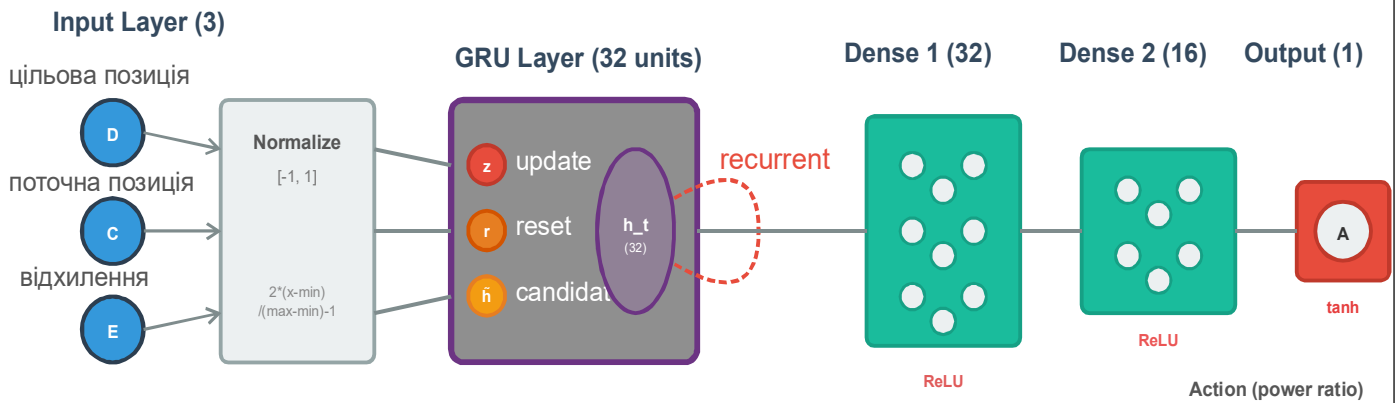
Через в'язке тертя (B) процеси проходять плавно, без коливань.

Рівняння динаміки :

$$\text{Динаміка: } T \frac{d\omega}{dt} = \omega - k \cdot U \quad \text{Кінематика: } \frac{d\phi}{dt} = \omega$$

					КРМ.АКСм-06.00.00.001			
Зм.	Арк.	№докум.	Підпис	Дата	Об'єкт керування і його математична модель	Літера	Маса	Мірило
Розроб.		Гулич						
Перев.		Стрілецький						
Т. контр.						Аркуш	Аркушів	
Н. контр.								
Затв.		Заміховський						

GRU Policy Network Architecture



					KPM.AKCM-06.00.00.002						
						Літера	Маса	Мірло			
Зм.	Арк.	№докум.	Підпис	Дата	Архітектура розробленої нейромережі для керування поворотною платформою						
Розроб.	Гулич										
Перев.	Стрілецький										
Т. контр.									Аркуш	Аркушів	
Н. контр.											
Затв.	Заміховський										

Процес навчання Evolution Strategies

1. Початкові випадкові ваги

$\theta_0 \sim$ random initialization
GRU Policy: ~6000 параметрів

2. Генерація шуму

$\epsilon \sim N(0, \sigma^2)$
Антитетичні пари: $+\epsilon$ та $-\epsilon$

3. Створення пертурбованих ваг (популяція)

$$\theta + \sigma \cdot \epsilon_1$$

$$\theta + \sigma \cdot \epsilon_2$$

...

$$\theta - \sigma \cdot \epsilon_1$$

$$\theta - \sigma \cdot \epsilon_2$$

...

4. Модель об'єкту (Table Simulator)

Кожна вага \rightarrow Епізод взаємодії з симулятором
Фізика: $J \cdot \omega = \tau - b \cdot \omega$
300 кроків $\times dt=0.1c$, випадкова ціль

5. Агенти оцінюють кожен крок епізоду

MoveToGoal
Покращення помилки

$$R_1(t)$$

TimeToGoal
Швидкість досягнення

$$R_2(t)$$

StayInThreshold
Утримання біля цілі

$$R_3(t)$$

SmoothAction
Плавність керування

$$-P_1(t)$$

Saturation
Уникнення насичення

$$-P_2(t)$$

На кожному кроці t : $Fitness(t) = R_1(t) + R_2(t) + R_3(t) - P_1(t) - P_2(t)$

6. Загальний фітнес для кожної ваги

$F = \sum Fitness(t)$ по всіх кроках епізоду

7. Пошук максимуму (градієнт)

Rank normalization фітнесів
 $\nabla \theta \approx \sum (F^* - F^-) \cdot \epsilon / (n \cdot \sigma)$

8. Мутація (Adam Optimizer)

$$m \leftarrow \beta_1 \cdot m + (1 - \beta_1) \cdot \nabla \theta$$

$$v \leftarrow \beta_2 \cdot v + (1 - \beta_2) \cdot \nabla \theta^2$$

9. Оновлення ваг

$$\theta_{t+1} = \theta_t + \alpha \cdot m / (\sqrt{v} + \epsilon)$$

$$\alpha = learning_rate = 0.01$$

Умова зупинки:

- max_generations
- patience (без покращення)

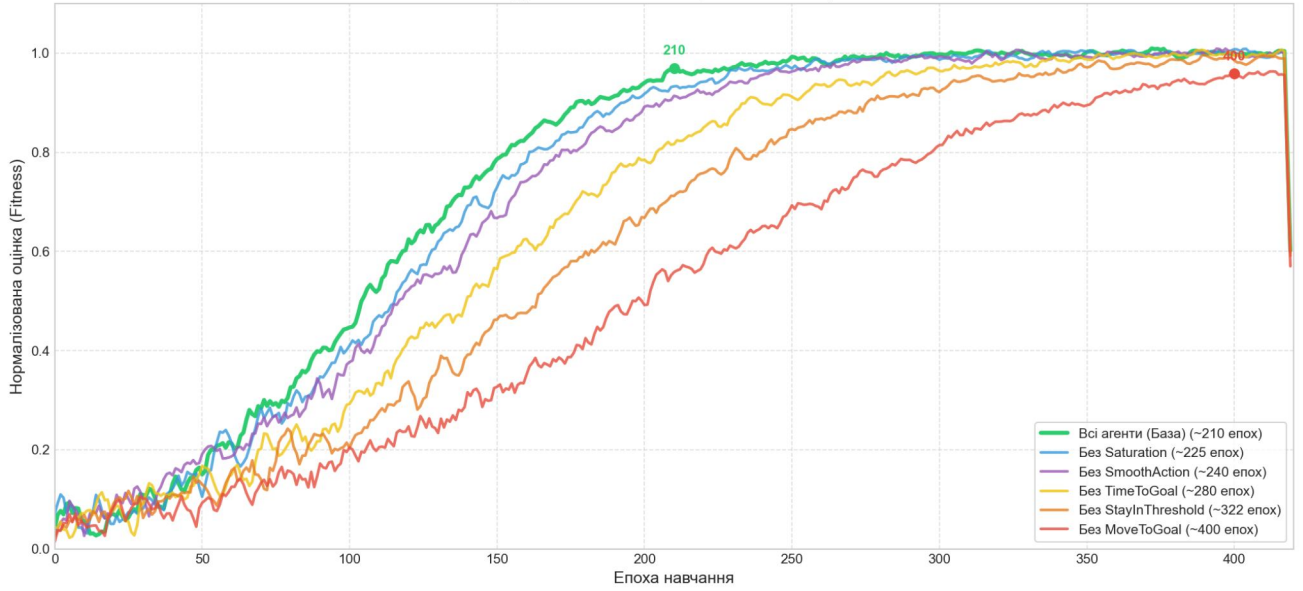
ПОВТОР ДОКИ НЕ ЗНАЙДЕНО МІНІМУМ

Це одна гілка обчислення для однієї популяції

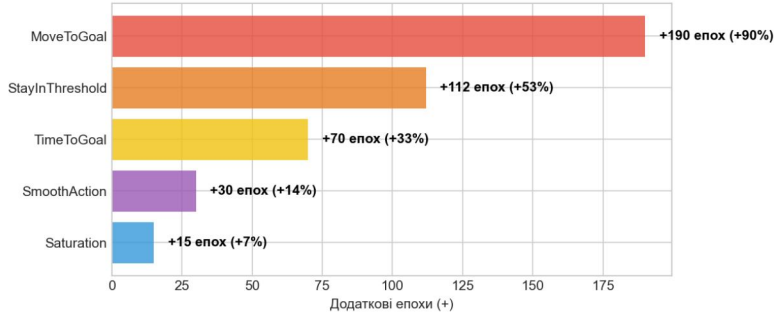
КРМ.АКСм-06.00.00.003

Зм.	Арк.	№докум.	Підпис	Дата	Реалізація процесу навчання розробленої нейромережі	Літера	Маса	Мірло
Розроб.		Гулич						
Перев.		Стрілецький						
Т. контр.						Аркуш	Аркушів	
Н. контр.								
Затв.		Заміховський						

Динаміка навчання (Fitness Score)



Вартість втрати агента (в епохах)

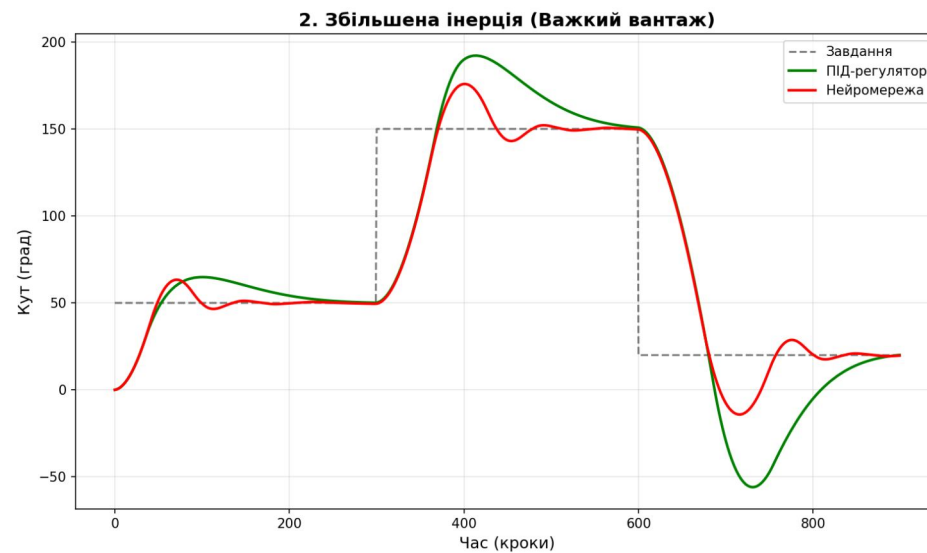
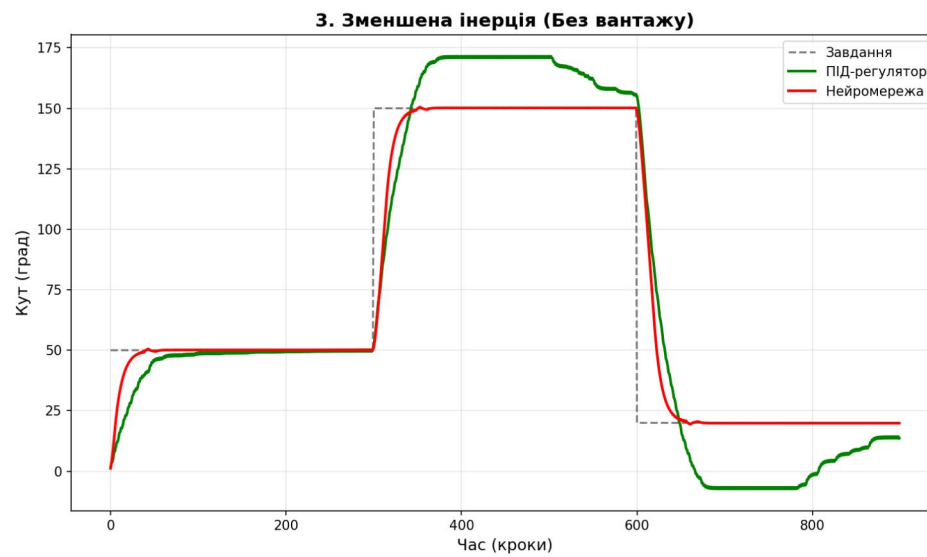
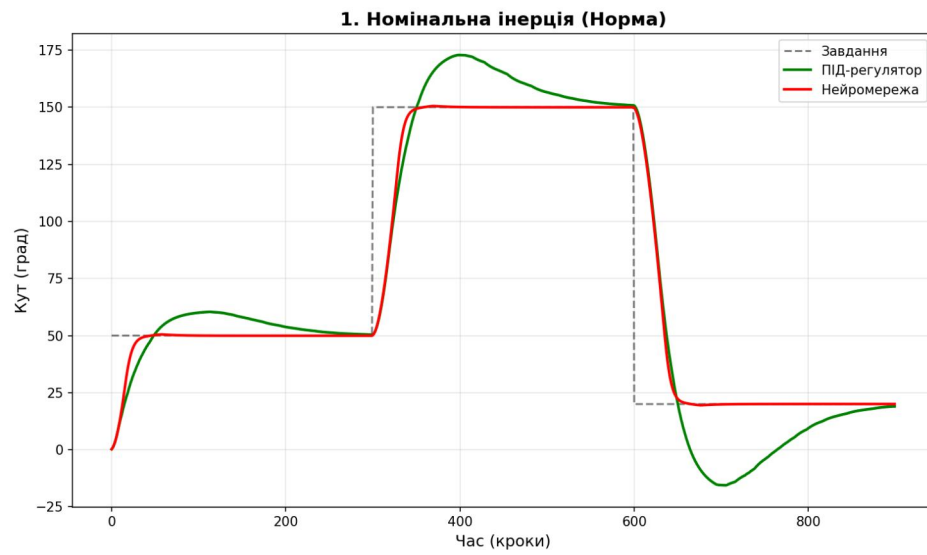


Класифікація важливості агентів

Агент	Епохи	Сповільнення	Важливість
Всі агенти	210	База	---
- Saturation	225	+7%	Низька
- SmoothAction	240	+14%	Низька
- TimeToGoal	280	+33%	Середня
- StayInThreshold	322	+53%	Висока
- MoveToGoal	400	+90%	Критична

КРМ.АКСм-06.00.00.004

Зм. Арк.	№докум.	Підпис	Дата	Аналіз ефективності агентів навчання	Літера	Маса	Мірло	
Розроб.	Гулич							
Перев.	Стрілецький							
Т. контр.					Аркуш	Аркушів		
Н. контр.								
Затв.	Заміховський							



KPM.AКСм-06.00.00.005

Зм.	Арк.	№докум.	Підпис	Дата
Розроб.		Гулич		
Перев.		Стрілецький		
Т. контр.				
Н. контр.				
Затв.		Заміховський		

Порівняльний аналіз ефективності застосування нейромережевого контролера

Літера	Маса	Мірло
Аркуш	Аркушів	