

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»
ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ СИСТЕМНОГО АНАЛІЗУ

ДИПЛОМНА РОБОТА НА ТЕМУ:

МОДУЛЬ ГЛИБИННОГО НАВЧАННЯ З ПІДКРІПЛЕННЯМ В СИСТЕМІ
АВТОНОМНОГО КЕРУВАННЯ АВТОТРАНСПОРТОМ

ВИКОНАВ СТУДЕНТ ГРУПИ КА-41

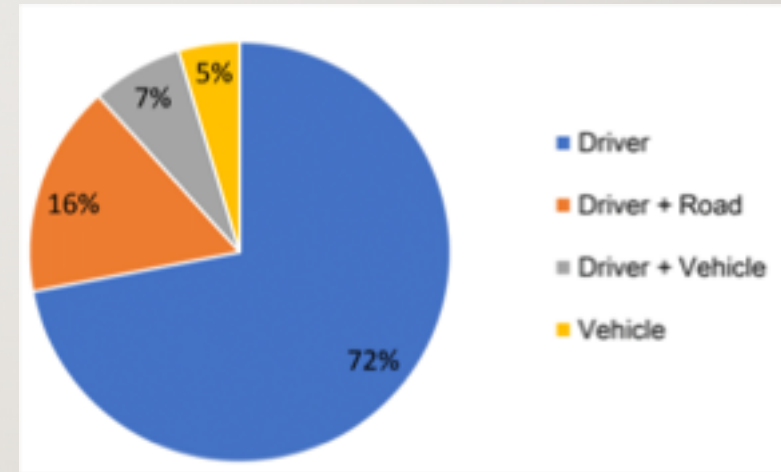
ТОКОВЕНКО Н.О.

НАУКОВИЙ КЕРІВНИК:

К.Т.Н ДОЦ. ДІДКОВСЬКА М.В

АКТУАЛЬНІСТЬ РОБОТИ

- Щорічно внаслідок ДТП гинуть близько 1.24 мільйона людей.
- За відсутності дій, згідно з прогнозами Всесвітньої організації охорони здоров'я, на момент 2030 року дорожньо-транспортні пригоди займуть 7 місце серед причин загибелі людей.



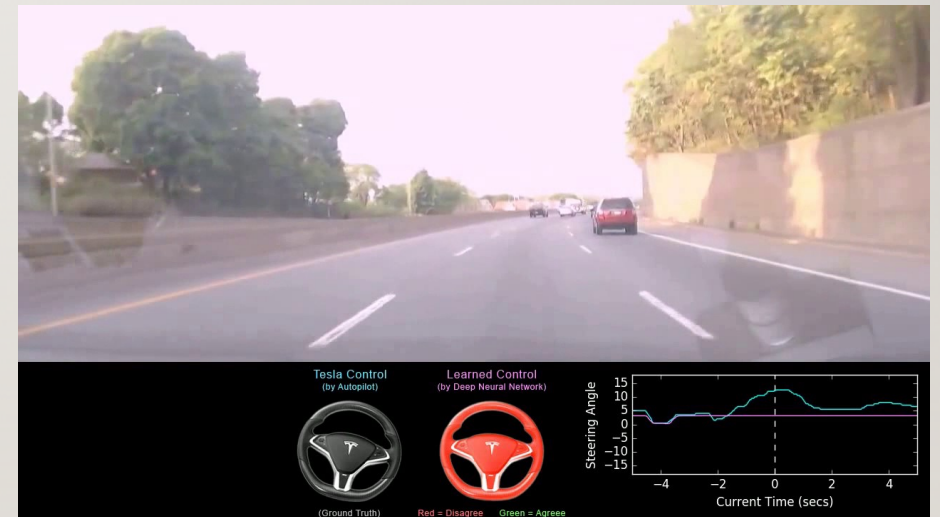
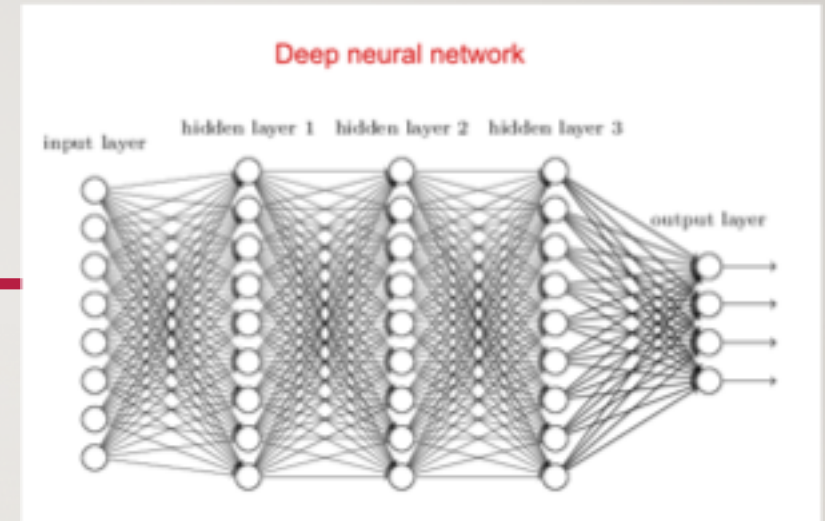
Причини ДТП

ПОСТАНОВКА ЗАДАЧІ

- Дослідження сучасних математичних моделей побудови інтелектуальних агентів для роботи в середовищі дорожнього руху
- Вивчення методик побудови архітектури систем існуючих програмних продуктів
- Реалізація різних підходів, проведення аналізу якості результатів
- Створення прототипу на основі отриманих результатів для подальших досліджень в даній сфері

ІСНУЮЧІ ПІДХОДИ

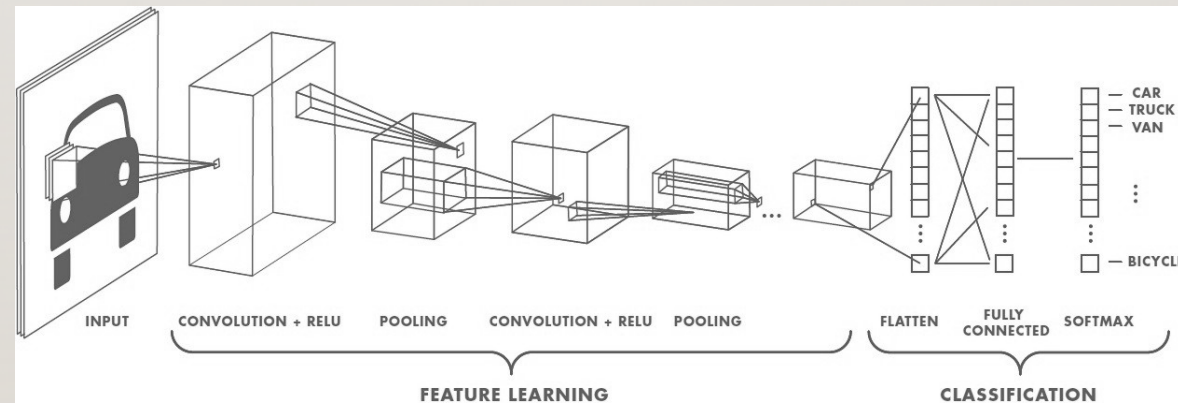
- Базуються на технології Supervised Learning з використанням глибоких нейронних мереж
- Переваги:
 - Відсутність необхідності виявлення ознак
 - Висока якість результатів прогнозу
- Недоліки:
 - Необхідність великої кількості даних для навчання
 - Неспроможність якісно діяти у ситуаціях, не передбачених розробниками
- Даний підхід вже використовується у комерційних продуктах флагманів даної сфери: Tesla, Google Waymo, Volvo



МАТЕМАТИЧНІ ОСНОВИ РІШЕНЬ

- Convolutional Neural Networks - це клас глибоких штучних нейронних мереж прямого поширення, який успішно застосовується до аналізу візуальних зображень.
- Головна перевага: використання згорткових шарів, що призводить до значного зменшення кількості параметрів моделі, що потребують навчання

$$s(t) = \int x(a)w(t - a)da = (x * w)(t)$$



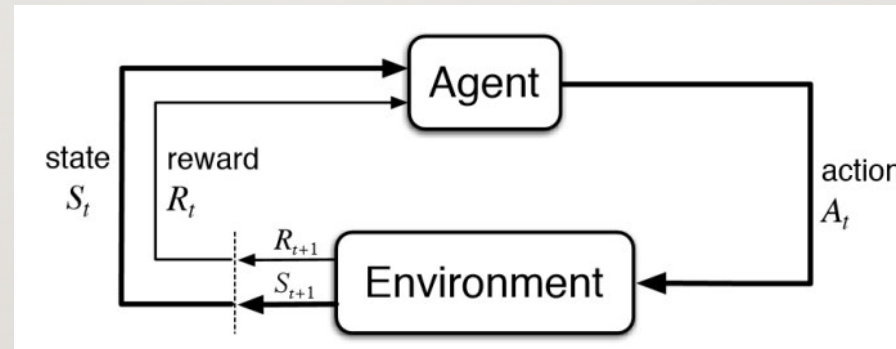
МАТЕМАТИЧНІ ОСНОВИ РІШЕНЬ

- Суть даних моделей: навчити параметри моделі таким чином, щоб прогнозовані значення навчальної вибірки мінімально відрізнялися від істинних.
- Навчання базується на використанні градієнтних методів оптимізації функції похибок:

$$L_{MSE}(\theta) = \frac{1}{N} \sum_{i=1}^N ((y_i - f(x_i, \theta))^2 \rightarrow \min$$

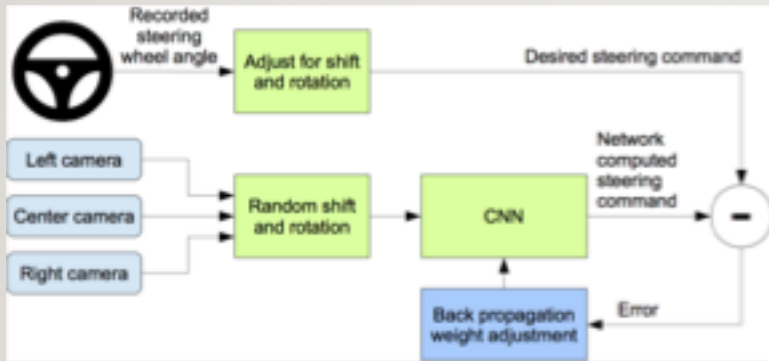
АЛЬТЕРНАТИВНІ ПІДХОДИ, ЩО РОЗГЛЯДАЮТЬСЯ В РОБОТІ

- Іншим потужним, але поки ще не дослідженим для вирішення даної задачі, класом алгоритмів машинного навчання є Reinforcement Learning.



- Задача агента полягає у тому, щоб діяти таким чином, щоб максимізувати власну винагороду

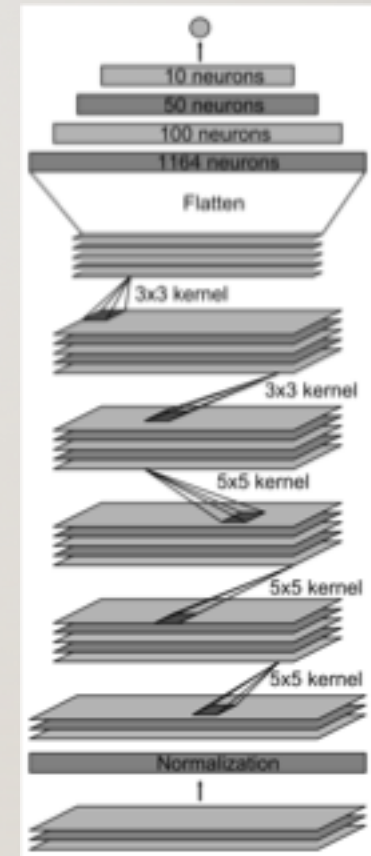
АРХІТЕКТУРИ АГЕНТІВ. АГЕНТ НА БАЗІ ВИКОРИСТАННЯ SUPERVISED LEARNING



- 11-шарова згорткова мережа.
- На вхід подається зображення з камери. На вихід прогнозоване значення повороту рульового керма.
- Всього 252219 параметрів
- Навчання за допомогою оптимізаційного методу Adam.
- Розподіл на навчальну і валідаційну вибірки: 80/20
- Розмірності даних при переході між шарами змінюються таким чином: $66 \times 200 \times 3 \rightarrow 31 \times 98 \times 24 \rightarrow 14 \times 47 \times 36 \rightarrow 5 \times 22 \times 48 \rightarrow 3 \times 20 \times 64 \rightarrow 1 \times 18 \times 64 \rightarrow 1164 \rightarrow 100 \rightarrow 50 \rightarrow 10 \rightarrow 1$
- Запобігання перенавчання моделі за допомогою техніки drop-out з ймовірністю вилучення клітини 0.5

АРХІТЕКТУРИ АГЕНТІВ. АГЕНТ НА БАЗІ ВИКОРИСТАННЯ SUPERVISED LEARNING

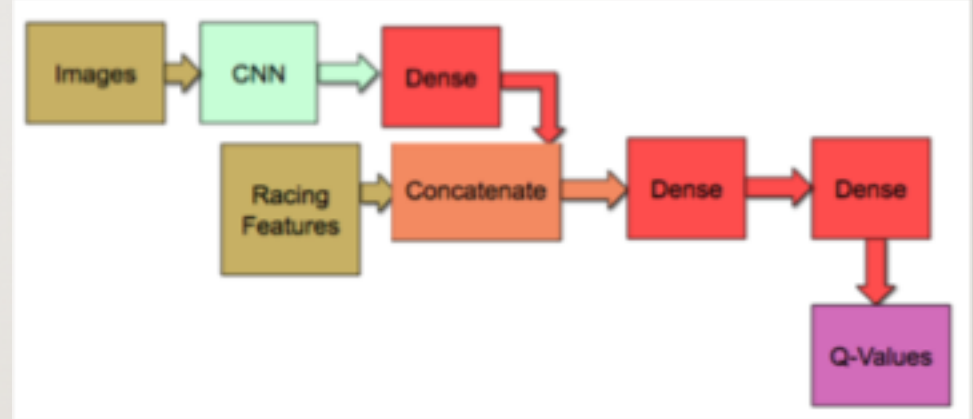
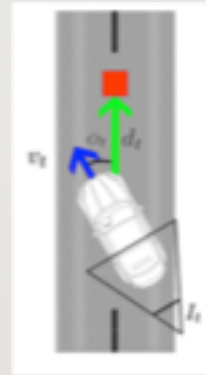
- Мотивація вибору такої архітектури нейронної мережі:
 - Емпірично виведено, що при початковій розмірності 66x200x3 слушним є використання 5 згорткових шарів. Саме шляхом стиску на кожному шарі у ~ 2 рази вдається отримати якісні ознаки з вхідного зображення
 - Внаслідок експериментів виявилось, що при використанні 4 повнозв'язних шарів вдається отримати найкращий прогноз. При збільшенні до 5 модель перенавчалась і давала гірші результати на валідаційній вибірці, а з 3 навпаки – прогнози були гіршими і на навчальній, і на валідаційній вибірках



АРХІТЕКТУРИ АГЕНТІВ. ГІБРИДНИЙ АГЕНТ НА БАЗІ ГЛИБИННОГО НАВЧАННЯ З ПІДКРІПЛЕННЯМ

- Використання динамічних характеристик автомобіля для опису стану агента
- $s_t = (I_t, v_t, d_t, \phi_t)$
 - I_t – зображення з камери, розташованої на лобовому склі,
 - v_t – оцінка вектору швидкості автомобіля,
 - ϕ_t – оцінка кута відхилення направляючого вектора руху автомобіля від направляючого вектора кривизни дороги.
 - d_t – оцінка відстані від центра автомобіля до центру дороги.

- Контроль зусилля двигуна:
$$throttle = 1 - a_t^2 - \frac{v_t^2}{(maxV)^2},$$



АРХІТЕКТУРИ АГЕНТІВ. ГІБРИДНИЙ АГЕНТ НА БАЗІ ГЛИБИННОГО НАВЧАННЯ З ПІДКРІПЛЕННЯМ

- Алгоритм навчання з підкріпленням Q-learning
- Розроблено власну структуру функції винагороди

function *Reward*(s_t, a_t):

$s_t = (l_t, v_t, d_t, \phi_t)$ – поточний стан агента

a_t – кут повороту рульового керма

$\alpha = 1.2$ – коефіцієнт підсилення

Порахувати час, необхідний для досягнення центру дороги: $\tau = \frac{d_t}{v_t \sin a_t}$

Обчислити фронтальне зміщення: $bias = v_t \cos a_t \tau$

Оцінити d_{t+1} : $\hat{d}_{t+1} = \left(\frac{v_t \tau}{\cos \phi_t} \right)^2 - bias^2$

$reward = e^{-\hat{d}_{t+1} \cdot \alpha}$

return *reward*

Запропонований критерій ефективності

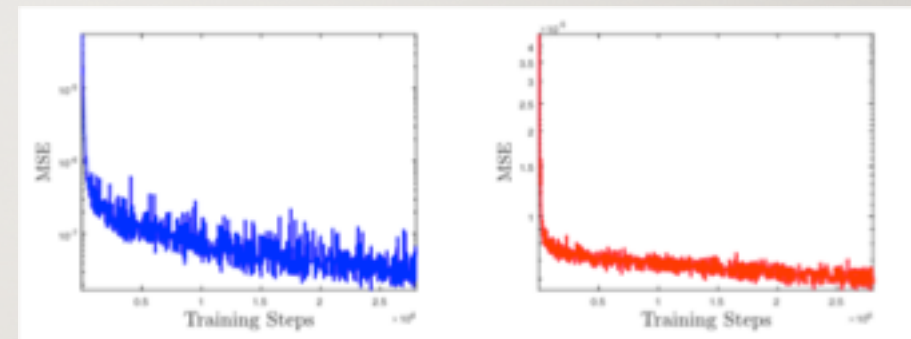
```
Initialize  $Q(s, a), \forall s \in S, a \in A(s)$ , arbitrarily, and  $Q(\text{terminal-state}, \cdot) = 0$ 
Repeat (for each episode):
  Initialize  $S$ 
  Repeat (for each step of episode):
    Choose  $A$  from  $S$  using policy derived from  $Q$  (e.g.,  $\epsilon$ -greedy)
    Take action  $A$ , observe  $R, S'$ 
     $Q(S, A) \leftarrow Q(S, A) + \alpha [R + \gamma \max_a Q(S', a) - Q(S, A)]$ 
     $S \leftarrow S'$ 
  until  $S$  is terminal
```

Псевдокод алгоритму Q-learning

АНАЛІЗ РЕЗУЛЬТАТІВ. ПОРІВНЯННЯ АРХІТЕКТУР

	Середня винагорода	Середня швидкість	Максимальна швидкість
Агент на базі SL	0.442	20.47 mph	24.91 mph
Гібридний агент (SL + RL)	0.567	13.57 mph	18.71 mph

Порівняльний аналіз агентів згідно впроваджених критеріїв ефективності



Криві навчання для навчальної та валідаційної вибірок

ДЕМОНСТРАЦІЯ РОБОТИ ПРОДУКТУ



Робота агента на базі SL



Робота гібридного агента на базі SL+RL

АНАЛІЗ РЕЗУЛЬТАТІВ

- Мотивація впровадження алгоритмів навчання з підкріпленням до систем автономного водіння полягає у їхній здатності приймати оптимальні рішення відносно заданих евристик у будь-якій ситуації, адже показники їхньої оптимальності залежать від доцільно визначених станів та обґрунтованого вибору можливих дій.
- Для ще більшого покращення роботи слід випробовувати більш довершені алгоритми, такі як Actor-Critic методи та ін.
- Головний результат полягає в тому, що гібридизація поведінки агентів здатна привести до більш стійких результатів. Саме тому можна вважати що на даному етапі дослідження проведено успішно.

ВИСНОВКИ

- Доведено доцільність залучення альтернативних підходів на базі глибинного навчання з підкріпленням до даної предметної області
- Запропоновано гібридну архітектуру побудови агентів на базі Reinforcement Learning
- Введено власний критерій ефективності у вигляді функції винагороди, згідно з яким можна будувати працюючі системи прийняття рішень в середовищі дорожнього руху
- Розроблене програмне забезпечення здатне як до власного використання, так і для використання пересічними користувачами, для подальших досліджень даної проблеми.

ШЛЯХИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

- Задача побудови автопілоту налічує ще безліч невирішених питань:
 - Планування подорожі
 - Дія в різних погодних умовах
 - Взаємодія з іншими учасниками дорожнього руху (пішоходи, інші машини) та об'єктами на дорозі
 - Дотримання правил дорожнього руху
 - Розпізнавання дорожніх контролюючих інструментів (світлофори, дорожні знаки)
 - Етична та законодавча підготовленість суспільства

ДЯКУЮ ЗА УВАГУ!