

# ФУНКЦІЇ АКТИВАЦІЇ В АРХІТЕКТУРІ НЕЙРОННИХ МЕРЕЖ

---

Виконав студент 1 року навчання, МП ПМ

Мокрий М. В.

Науковий керівник

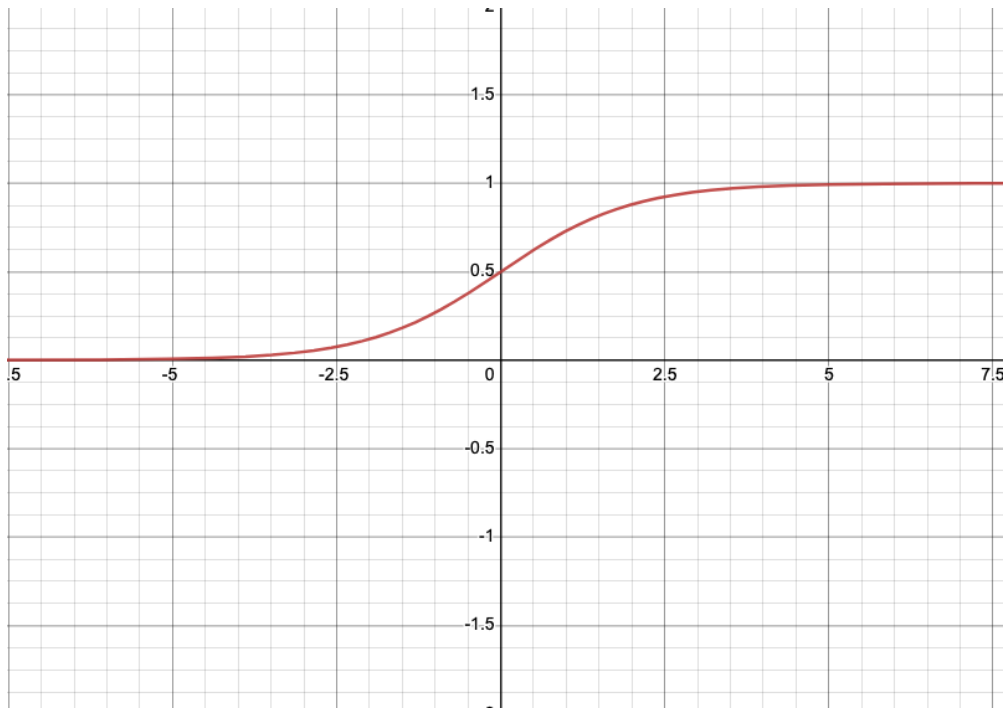
ст. викладач Швай Н. О.

# Основними завданнями даного дослідження є:

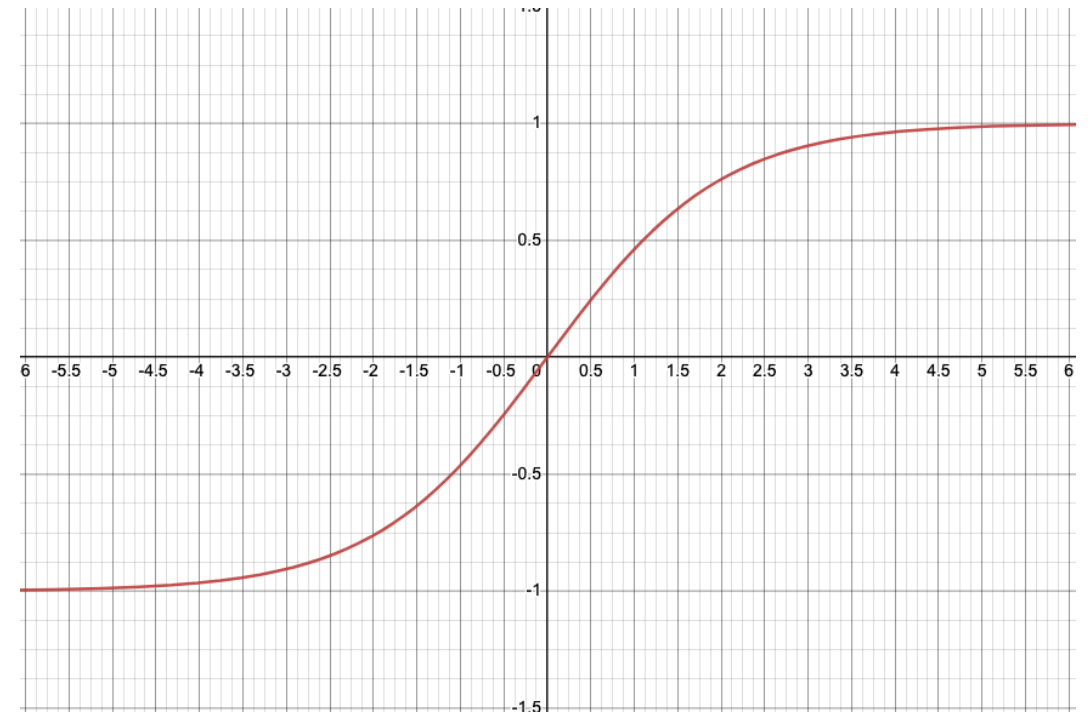
---

- Опис різних функцій активації, їх особливостей та проблем;
- Дослідження роботи функцій активації на різних моделях та архітектурах штучних нейронних мереж;
- Аналіз кінцевих результатів та пошук закономірностей між ними;
- Визначення від яких параметрів залежить вибір функції активації.

# Функції активації нейронних мереж



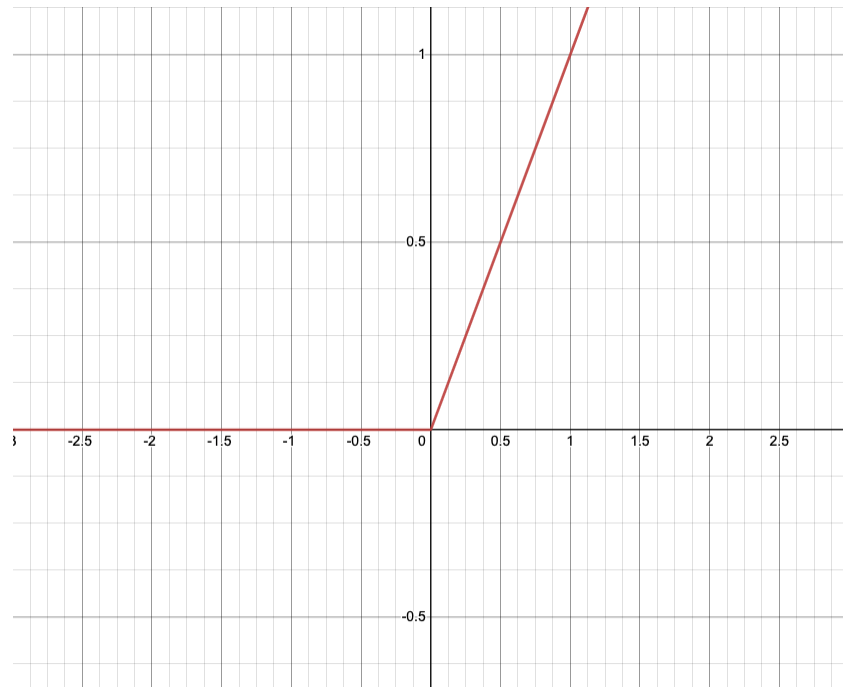
Сигмоїдна функція



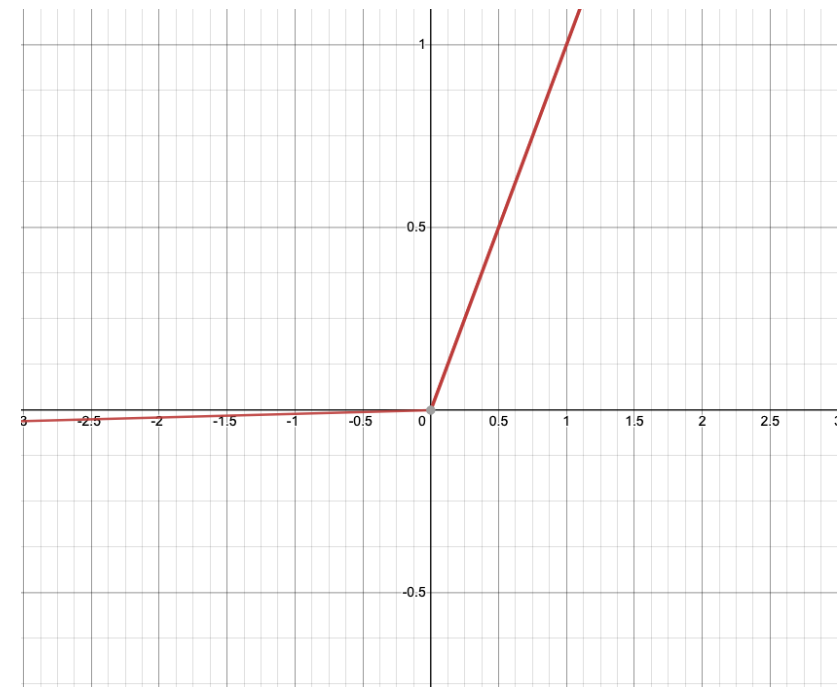
Тангенсна функція

# Функції активації нейронних мереж

---



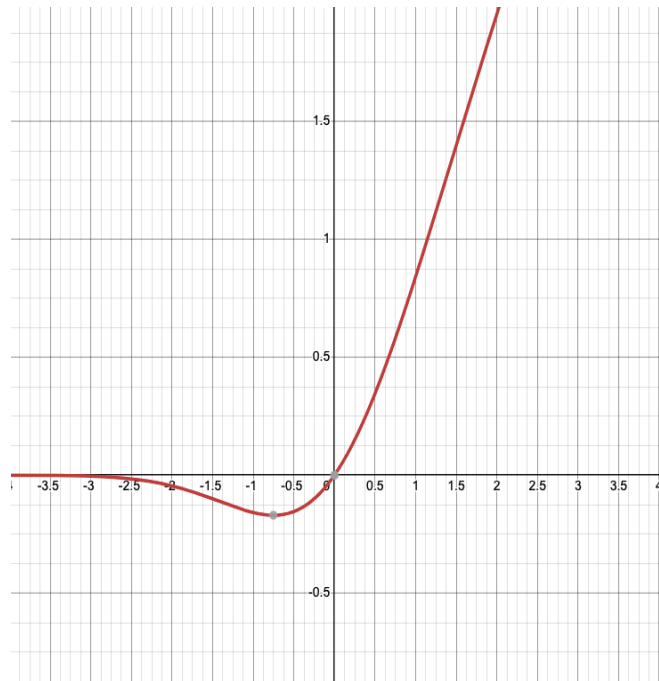
ReLU



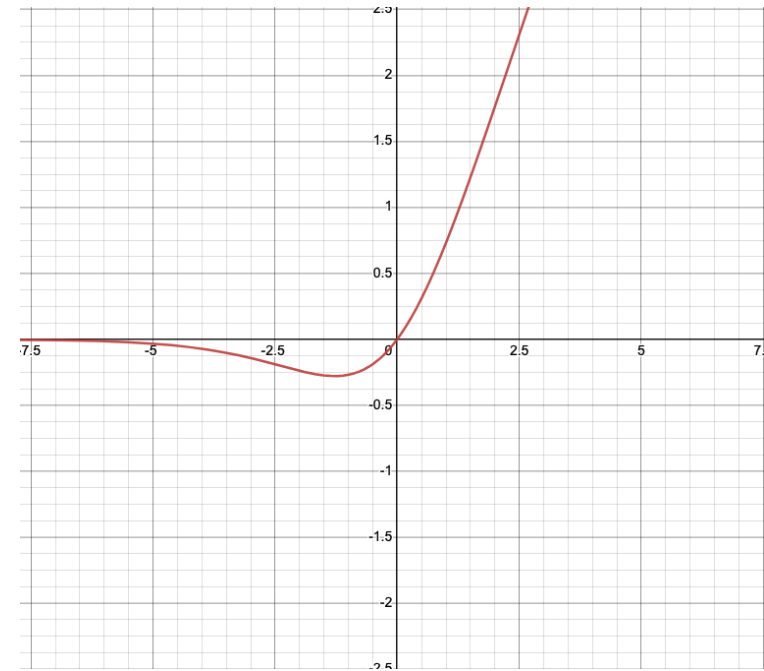
LReLU

# Функції активації нейронних мереж

---



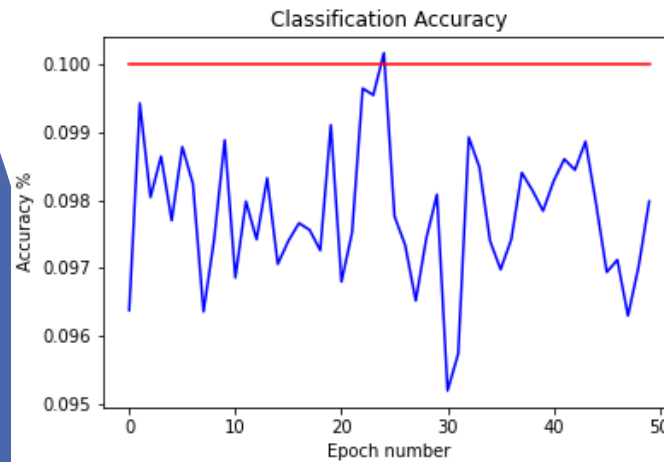
GeLU



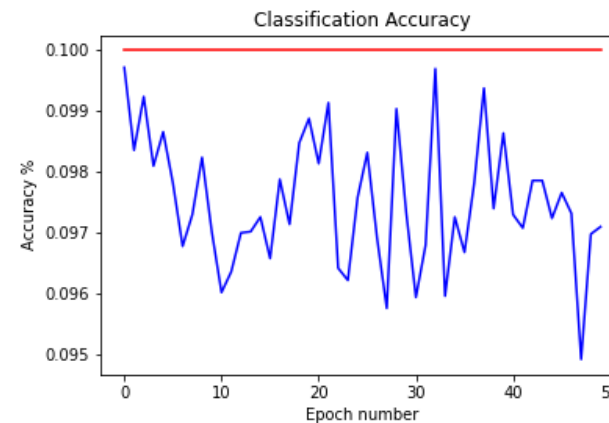
Swish

# Сигмоїдна функція

## VGG

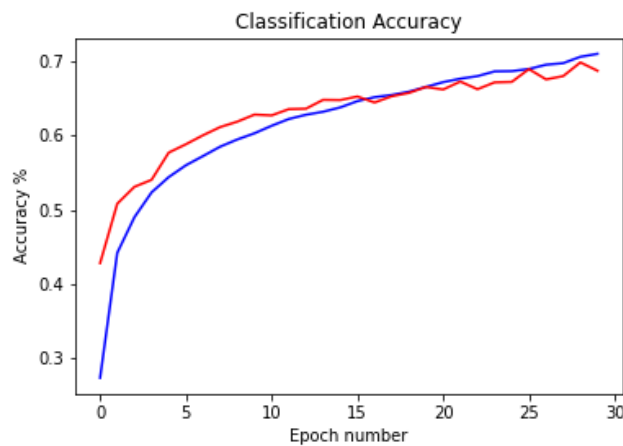


He

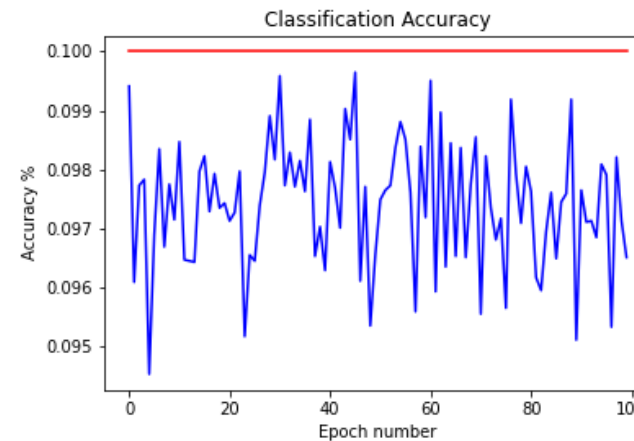


Xavier

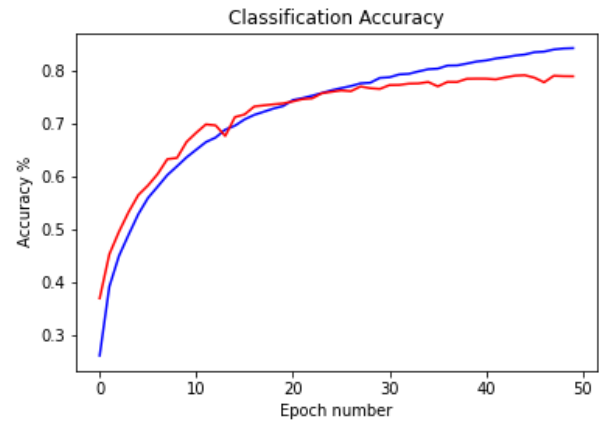
## ViT



## ResNet

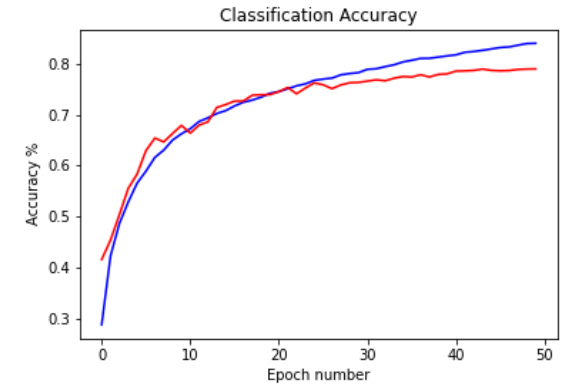


# Тангенсна функція



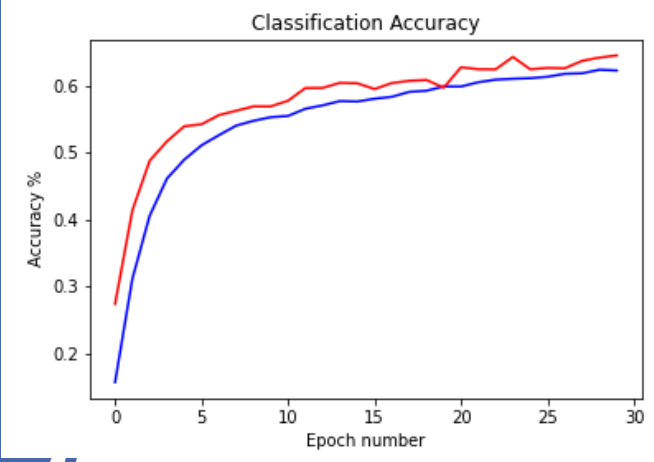
He

# VGG

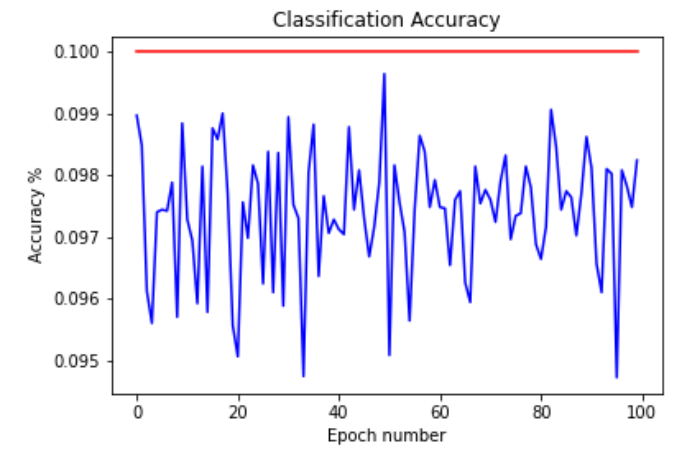


Xavier

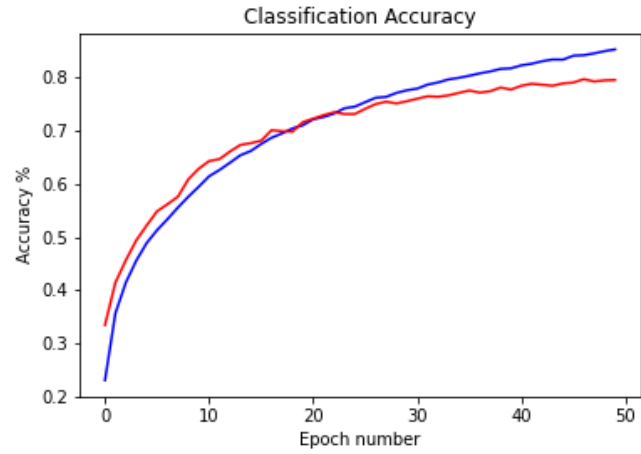
# ViT



# ResNet

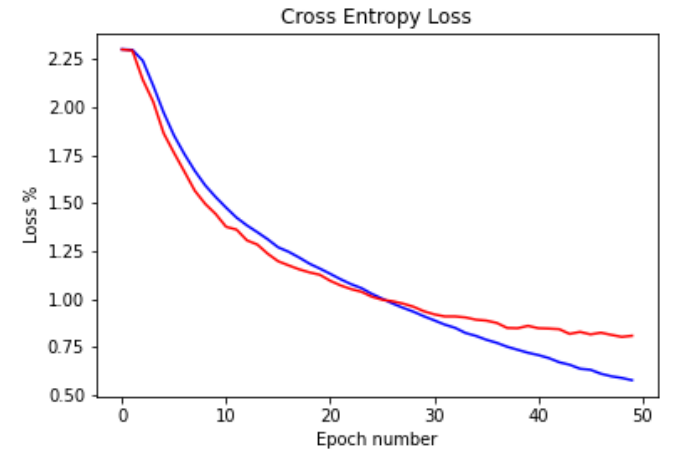


Swish



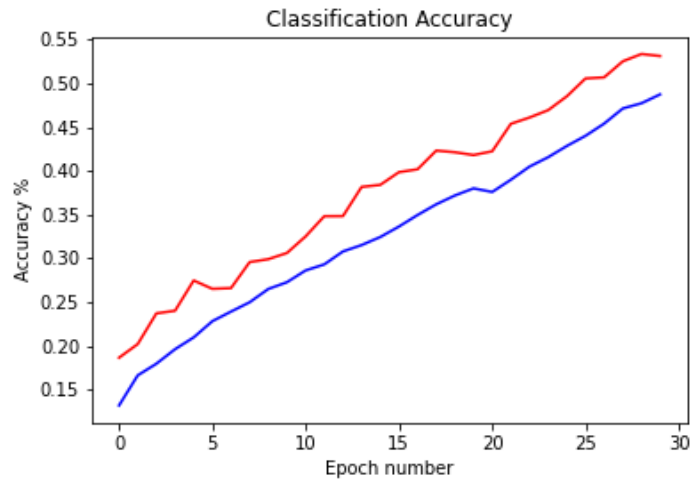
He

VGG

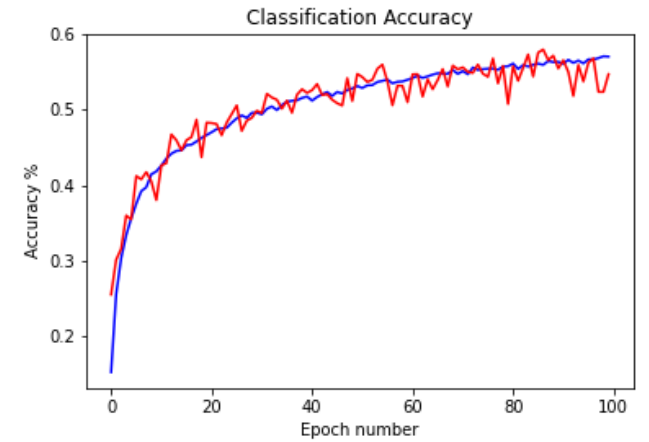


Xavier

ViT

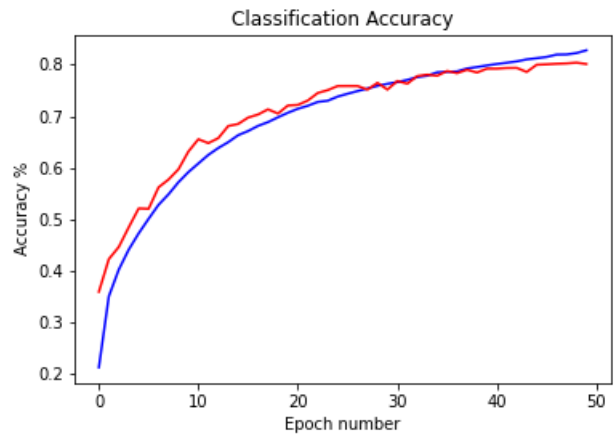


ResNet



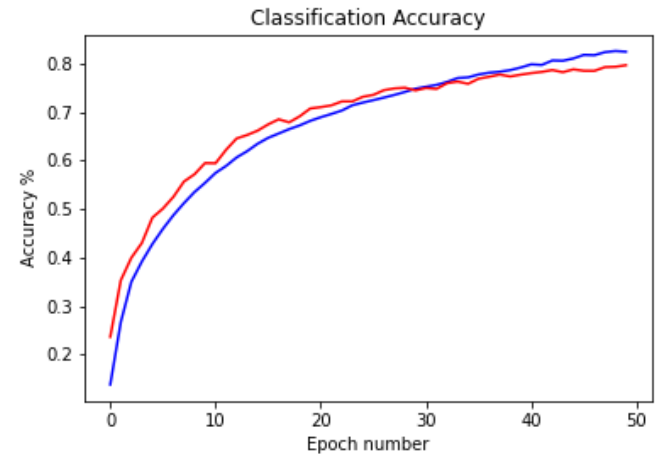


ReLU



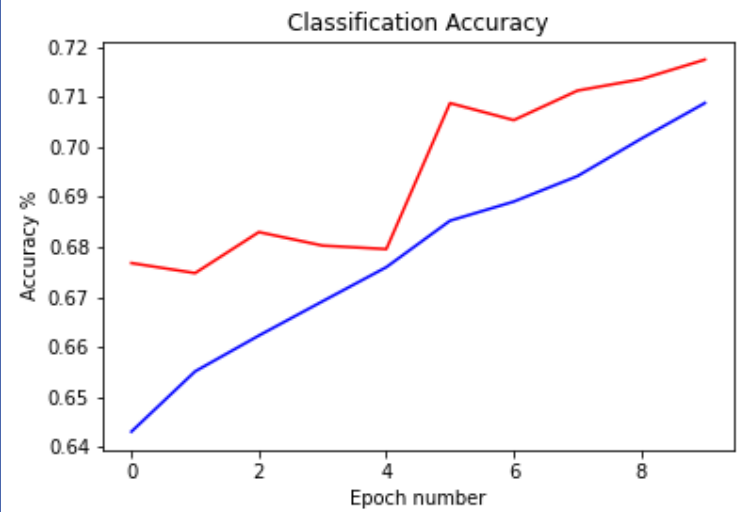
He

VGG

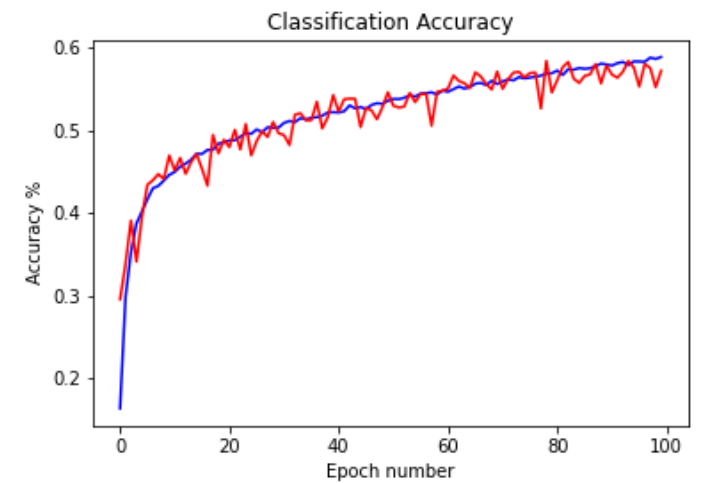


Xavier

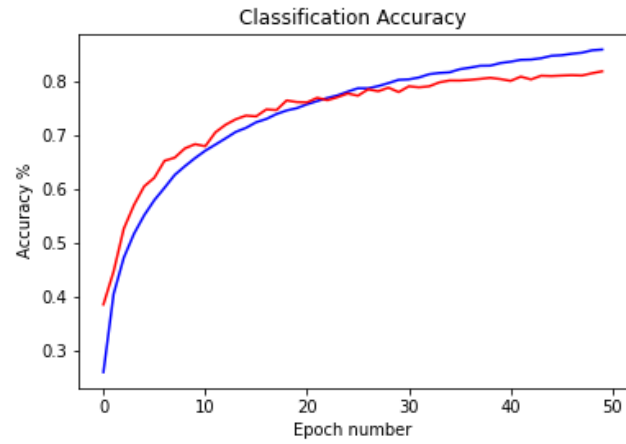
ViT



ResNet

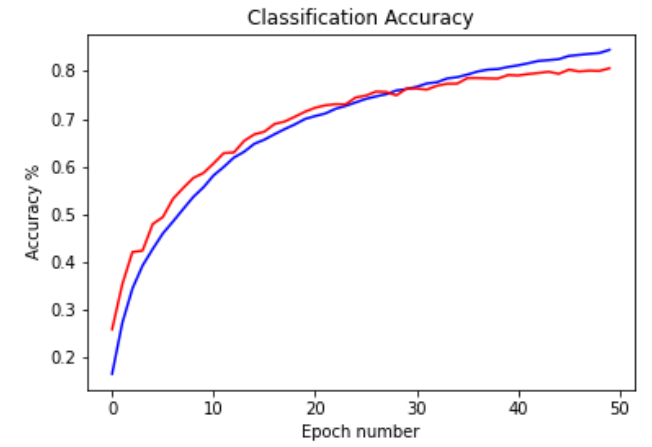


LReLU



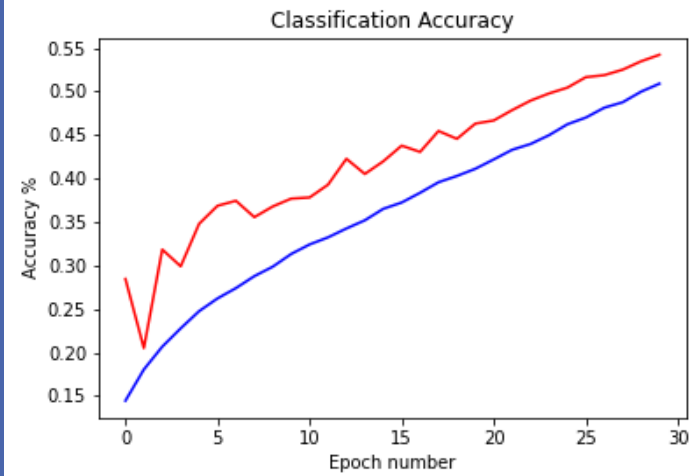
He

VGG

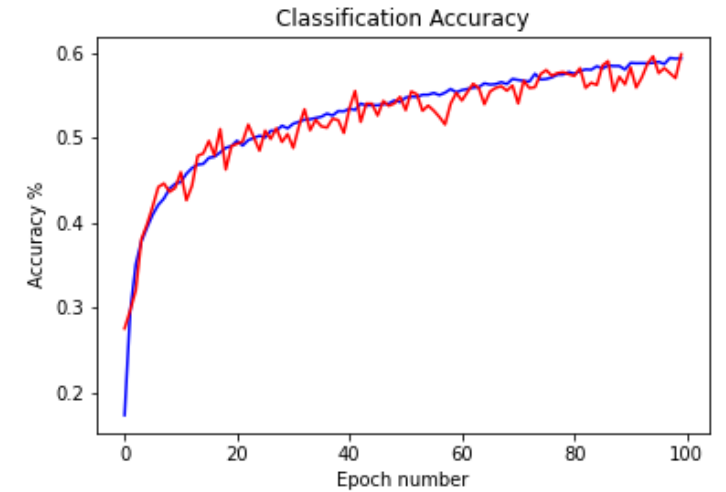


Xavier

ViT

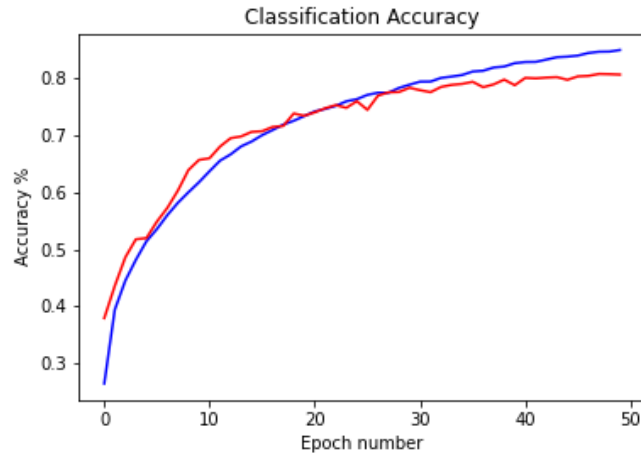


ResNet

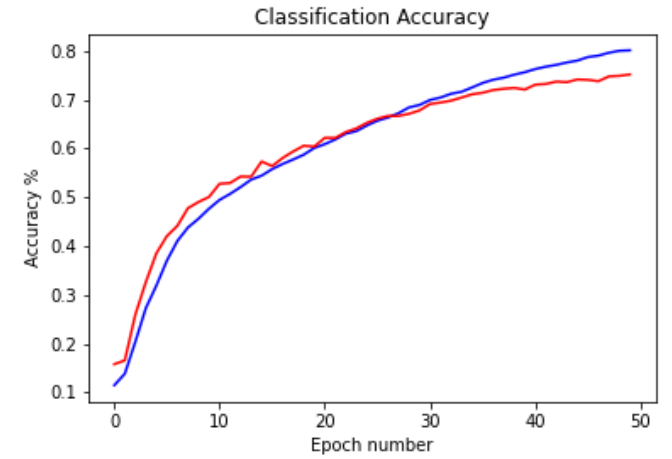


GeLU

VGG

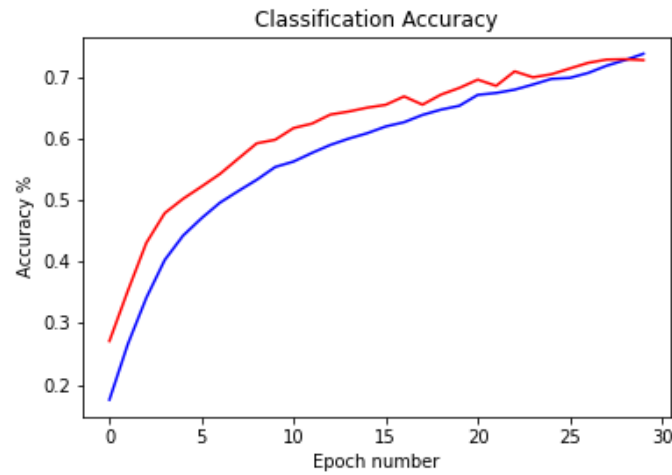


He

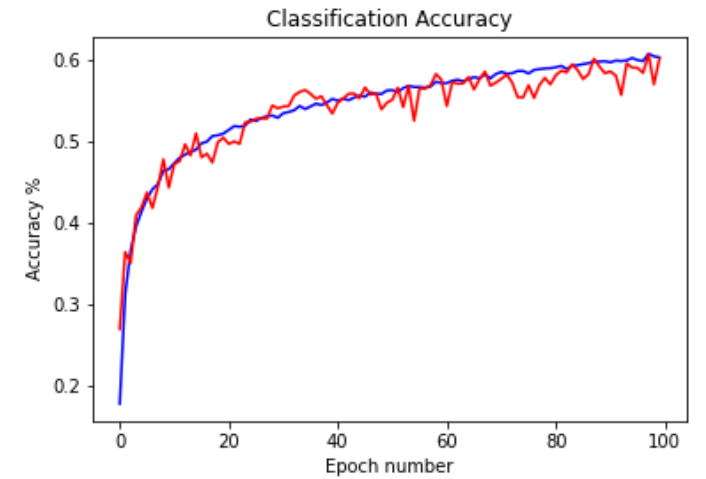


Xavier

ViT



ResNet



# Результати роботи

Функція активації	VGG (He) Accuracy	VGG (Xavier) Accuracy	ViT Accuracy	ResNet Accuracy
Sigmoid	0.1000	0.1000	0.6872	0.1000
Tanh	0.7895	0.7903	0.6289	0.1000
Swish	0.7956	0.7273	0.5212	0.5473
ReLU	0.8015	0.7965	0.7174	0.5716
LReLU	0.8184	0.8060	0.5421	0.5980
GeLU	0.8066	0.7520	0.7290	0.6018

# Висновки

---

- Результатом виконаної роботи став огляд найбільш популярних функцій активації штучних нейронних мереж та аналіз їх ефективності на різних архітектурах нейронних мереж з різними методами ініціалізації ваг.
- Було досліджено, що функції активації напряму залежать від архітектури штучної нейронної мережі, її конкретної моделі, а також метода ініціалізації ваг.
- У ході проведеного експерименту було визначено, що найбільш стабільною функцією активації залишається ReLU, яка на всіх моделях показала хороший результат, а найбільш спірною є функція активації сигмоїд, яка аж на двох моделях не змогла допомогти натренувати модель. В той час найбільш результативною стала функція активації GeLU, показавши на двох різних моделях найкращий результат.

Дякую за увагу!

---