

**Deep Learning BootCamp**

**?**

**Hackathon**

# Deep Learning BootCamp



With Subrat Panda and Michael Yugas



# Agenda

Time	Item
2:00-2:05	Challenge Description and Dataset
2:05-2:10	Contest Rules
2:10-2:30	Reference Implementation
2:30-5:00	Q + A



?

POKÉMON

# Dataset

- Guess a Pokémon's primary type and secondary type based on an image and a textual description
- Only Gen I Pokémon
- The training set does not contain all 151 Pokémon; your algorithm needs to work even on unseen data
- This is a *multimodal* classification problem
- Test set is blind: you do not get to see it

## Training Set Stats

- 6175 Samples
- 19 Primary Types
- 19 Secondary Types
- 136 Pokémon
- Average Image Size: 723x723
- Average Description Length: 20 words

# Example Data

Image:



Description:

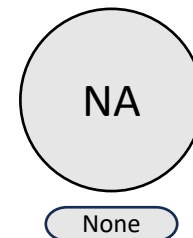
“It swims facing backward by opening and closing its two-piece shell. It is surprisingly fast.”

**Model Input (x)**

Primary Type:



Secondary Type:



**Targets (y)**





<https://github.com/ntu-dl-bootcamp/deep-learning-2024/blob/main/hackathon/hackathon.ipynb>

**Scan to Download**

# Rules

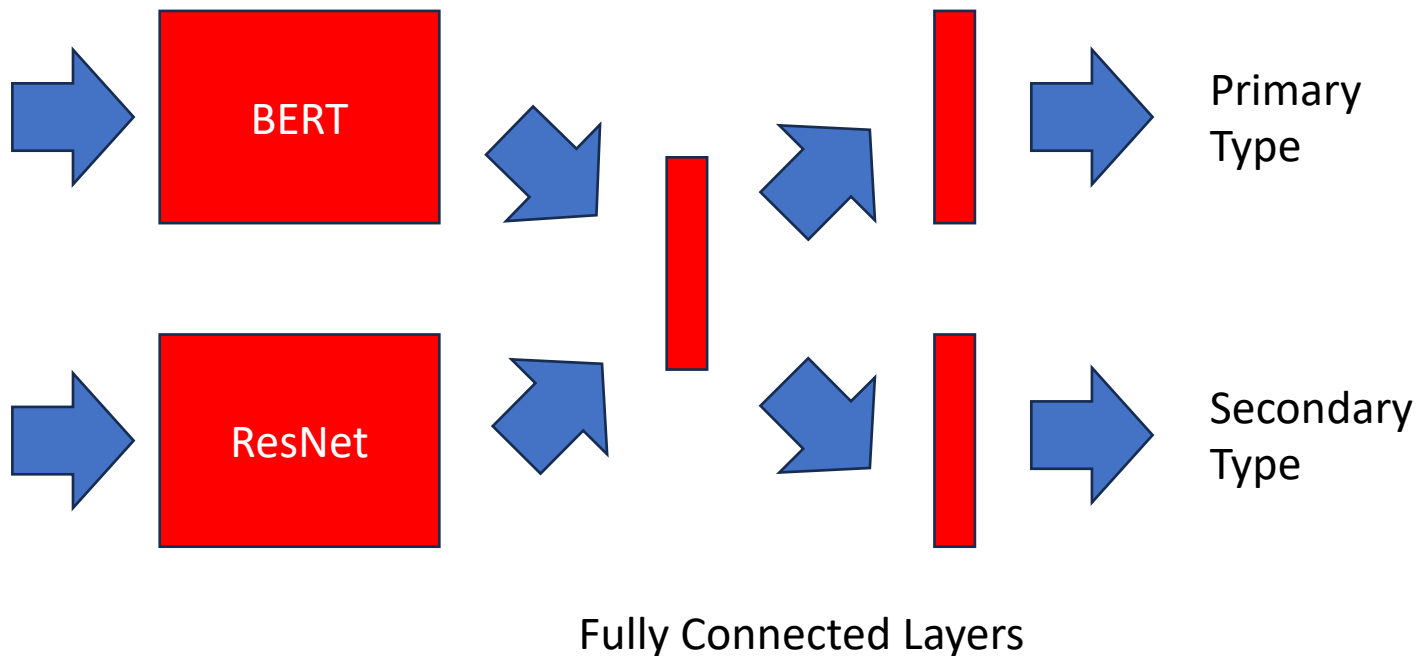
- You can only use the dataset provided to you. Hardcoding a table of Pokémon names and their corresponding types is forbidden.
- You have until the 23:59 on Saturday (30-Mar-2024) to complete the challenge
- You are allowed to use the Internet, ChatGPT, or any other resources to help you write your code, but you must train your model yourself
- You are allowed to start with a pretrained model, including transformers.
- You must submit both your iPython notebook (\*.ipynb) and the weights (\*.pt) file by email
- The winner will be the team with the best accuracy on primary type classification, ties will be broken by performance on secondary type accuracy
- You cannot get help from other teams enrolled in the competition
- You cannot get help from student studying computer science or a related field if they are not involved in the competition and enrolled on your team



# Reference Implementation



“It swims facing backward by opening and closing its two-piece shell. It is surprisingly fast.”



- At a minimum you need:
  - Feature extraction
  - Mode combination
  - 2 classification heads

- Feel free to go off the beaten path:
  - How early do you combine features?
  - How deep are your heads?
  - How much processing on the combined feature space?



**Q+A**