# B351 Project

## Logits

- Logit represents the log odds of the event happening
  - $logit(p) = ln(\frac{p}{1-p})$
    - Where $p$ = the probability of the class
    - As $p \to 0,\ logit(p) \to -\infty$
    - As $p \to 1,\ logit(p) \to +\infty$
  - Used by logistic regression and classification networks
  - Softmax and sigmoid functions converts logits to probabilities
- In our code we use logits and then argmax as we use this specific huggingface model for testing, not training
  - Argmax shows the highest class, whereas softmax creates a probability distribution
    - For testing we only need one class, for training we need the probability of all classes
  - When we train our own model we will use logits for a softmax function

## Model #1

- "facial_emotions_image_detection" by dima806
  - Uses a base vision transformer "vit-base-patch16-224-in21k" by Google
- Fine tuned for the task of facial emotion recognition
  - 7 emotions: Sad, disgust, angry, neutral, fear, surprise, happy
- id2label mapping:
  - 0 → sad
  - 1 → disgust
  - 2 → angry
  - 3 → neutral
  - 4 → fear
  - 5 → surprise
  - 6 → happy
- Achieved 57% accuracy in personal testing, claimed to achieve 90% with original dataset.

## Dataset #1

- RAF-DB Dataset by Dev-Shuvoalok
- Contains 15,000 images in 100x100 pixels
  - 12,000 training sample and 3,000 test sample
- AutoImageProcessor auto resizes them to 224 × 224 for vision transformer
- Images in this database are of great variability in:
  - Subjects' age
  - Gender and ethnicity
  - Head poses
  - Lighting conditions
  - Occlusions, (e.g. glasses, facial hair or self-occlusion)
  - Post-processing operations (e.g. various filters and special effects)