**Lecture Notes – Loss Function for Classification (Cross-Entropy)**

**1. Regression vs. Classification**

* **Regression**: Output is a continuous number (e.g., price, height).
* **Classification**: Output is a category (e.g., cat, dog, horse).
* Need a different loss function for classification.

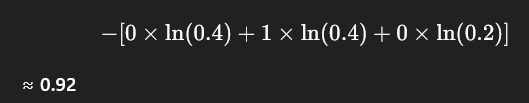
**2. Common Loss Function for Classification**

* **Cross-Entropy Loss**  
  Formula:
  + **Target (T)**: The correct category in vector form.
  + **Output (Y)**: Model’s predicted probabilities for each category.

**3. Example – Cats, Dogs, Horses**

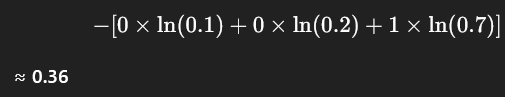
**Case 1: Image labeled as Dog**

* **Target vector (T)**: [0, 1, 0]  
  (0 → not cat, 1 → dog, 0 → not horse)
* **Model output (Y)**: [0.4, 0.4, 0.2]  
  (40% cat, 40% dog, 20% horse)
* **Cross-Entropy Calculation**:



**Case 2: Image labeled as Horse**

* **Target vector (T)**: [0, 0, 1]
* **Model output (Y)**: [0.1, 0.2, 0.7]  
  (10% cat, 20% dog, 70% horse)
* **Cross-Entropy Calculation**:



**4. Interpretation**

* **Lower cross-entropy** → Better prediction.
* In the example:
  + First image loss (0.92) is higher → Model uncertain (40% cat, 40% dog).
  + Second image loss (0.36) is lower → Model more confident (70% horse).

**5. Simplification for Classification**

* Target vectors have:
  + **One "1"** (correct category).
  + **Rest "0"s** (other categories).
* This means cross-entropy can be simplified to:

−ln⁡(Probability of the correct class)-\ln(\text{Probability of the correct class})−ln(Probability of the correct class)

**6. Final Notes**

* **Most common loss functions**:
  + Regression → **Squared Loss**.
  + Classification → **Cross-Entropy Loss**.
* **Any function** that:
  + Is **higher for worse results**,
  + **Lower for better results**,  
    can be used as a loss function.
* This flexibility is useful when **coding**.