PROBLEM 1

Q1)

$$o_i = \frac{exp(v_i)}{\sum_{j=1}^{k} exp(v_k)}$$

For

$$i \neq j,\ \frac{\partial o_i}{\partial v_j} = exp(v_i)\frac{\partial \frac{1}{\sum_{j=1}^{k} exp(v_k)}}{\partial v_j} = exp(v_i)\frac{-exp(v_j)}{(\sum_{j=1}^{k} exp(v_k))^2} = -\frac{exp(v_i)}{\sum_{j=1}^{k} exp(v_k)} * \frac{exp(v_j)}{\sum_{j=1}^{k} exp(v_{\cdot}}$$

For

$$i = j,\ \frac{\partial o_i}{\partial v_j} = \frac{exp(v_i)*[\sum_{j=1}^{k} exp(v_k) - exp(v_i)]}{(\sum_{j=1}^{k} exp(v_k))^2} = \frac{exp(v_i)}{\sum_{j=1}^{k} exp(v_k)} * \frac{\sum_{j=1}^{k} exp(v_k) - exp(v_i)}{\sum_{j=1}^{k} exp(v_k)} = o_i * (1$$

Q2)

$ \frac{\partial L}{\partial v_i} = \frac{\partial [-\sum_k y_k log(o_k)]}{\partial v_i} = -\sum_k \frac{\partial y_k log(o_k)}{\partial v_i}= -\sum_k y_k \frac{\partial log(o_k)}{\partial v_i} =-\sum_k y_k \frac{\partial log(o_k)}{\partial o_k}

\frac{\partial o_k}{\partial v_i}= -\sum_k y_k \frac{1}{o_k} \frac{\partial o_k}{\partial v_i}=$

$$\frac{\partial L}{\partial v_i} = [-\sum_{k \neq i} y_k \frac{1}{o_k}(-o_i o_k)] - y_i * \frac{1}{o_i} * (o_i(1 - o_i))$$ **(using Q1)**

$$\frac{\partial L}{\partial v_i} = [-\sum_{k \neq i} y_k(-o_i)] - y_i * (1 - o_i) = \sum_{k \neq i} y_k(o_i) + (y_i o_i) - y_i = o_i(\sum_k y_k) -$$

Also, $\sum_k y_k = 1$ because is the one-hot encoded class label and therefore,

$$\boxed{\frac{\partial L}{\partial v_i} = o_i - y_i}$$

In [ ]:

```

```