تمرین کوتاه 4 – شبکه های عصبی – مازندرانیان – 830402066

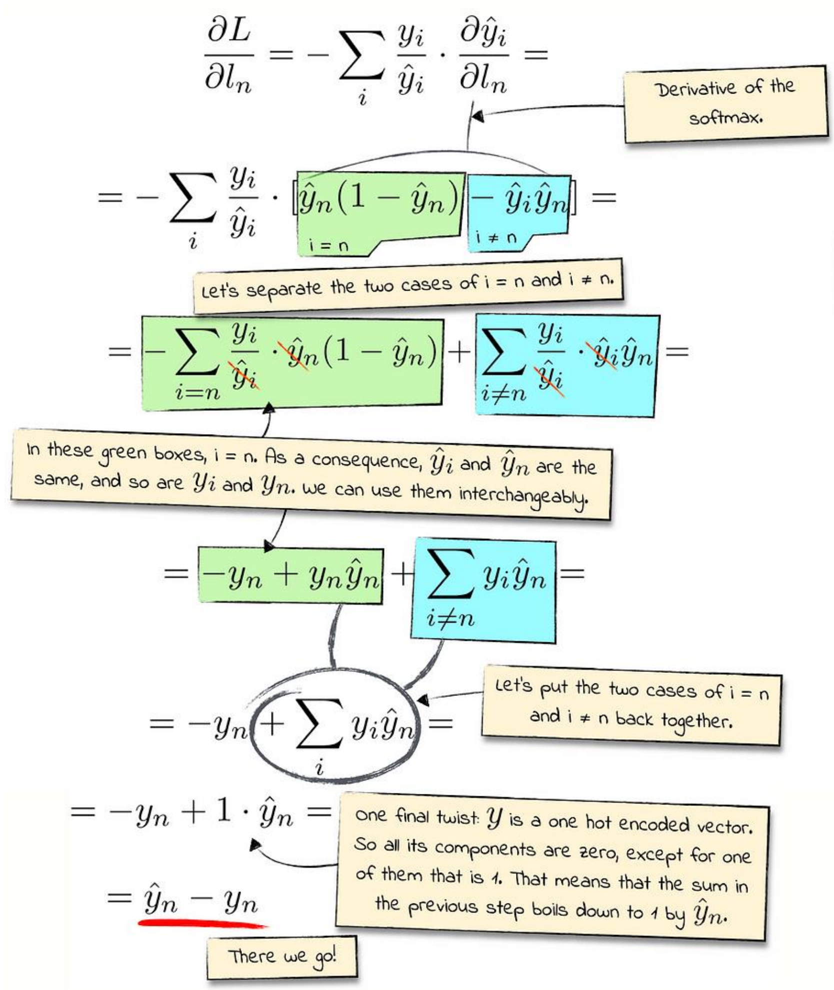
سوال 1) اثبات فرمول Log Loss در مسائل Multi-Class و استفاده از تابع SoftMax

انتروپی کراس (Cross-Entropy) بین دو توزیع احتمال p و q به صورت زیر تعریف می‌شود:

$$H(p,q) = -\Sigma(p(x) * log(q(x)))$$

فرض کنید یک نمونه با K کلاس داریم. برای این نمونه، مدل ما یک بردار احتمال $p = (p1, p2, ..., pK)$ تولید می‌کند که در آن pi احتمال تعلق نمونه به کلاس i است. همچنین، برچسب واقعی نمونه را با یک بردار یک‌کدی y نشان می‌دهیم که در آن عنصر مربوط به کلاس صحیح برابر با 1 و بقیه عناصر برابر با 0 هستند. در این صورت، Log Loss به صورت زیر محاسبه می‌شود:

$$L = -\Sigma(yi * log(pi))$$



سوال 2) تشخیص Convex بودن تابع با استفاده از مشتق دوم

فرض کنید تابع $f(x) = x^2$ را داریم. مشتق اول این تابع $f'(x) = 2x$ و مشتق دوم آن $f''(x) = 2$ است. از آنجایی که مشتق دوم همواره مثبت است، این تابع محدب است.

فرض کنید تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}$ دو بار پیوسته مشتق‌پذیر باشد. این تابع در دامنه خود محدب است اگر و تنها اگر ماتریس هسیان (ماتریس مشتقات جزئی مرتبه دوم) آن در هر نقطه از دامنه، نیمه‌مثبت معین باشد.

- **ماتریس هسیان :**ماتریسی است که درایه‌های آن مشتقات جزئی مرتبه دوم تابع هستند.
- **نیمه‌مثبت معین :**یک ماتریس A نیمه‌مثبت معین نامیده می‌شود اگر برای هر بردار غیرصفرx ، حاصلضرب* x^T $A * x$ بزرگتر یا مساوی صفر باشد.

در حالی که مثبت بودن مشتق دوم در تمام نقاط دامنه یک شرط **کافی** برای محدب بودن تابع است، اما شرط **ضروری** نیست. به عبارت دیگر، اگر تابعی محدب باشد، لزوماً مشتق دوم آن در همه نقاط مثبت نیست.

توابع غیرمشتق‌پذیر :بسیاری از توابع محدب وجود دارند که در برخی نقاط مشتق‌پذیر نیستند. برای مثال، تابع مقدار مطلق |x|در نقطه x=0 مشتق‌پذیر نیست، اما در کل یک تابع محدب است .

محدب بودن در نقاط عطف :در نقاط عطف یک تابع، مشتق دوم صفر می‌شود. این بدان معناست که در این نقاط، شرط مثبت بودن مشتق دوم نقض می‌شود، اما تابع ممکن است همچنان در اطراف این نقاط محدب باشد.