

به نام آنکه جان را فکرت آموخت



دانشگاه صنعتی شریف دانشکده مهندسی برق

آزمون پایانی درس یادگیری عمیق سال تحصیلی ۱۴۰۰-۰۱

نام و نام خانوادگی دانشجو:

شماره دانشجویی:

نکات مهم:

نکات مهم:

- آزمون شامل شش سوال و مسئله در سه صفحه (شامل این صفحه) و مدت آن ۱۲۰+۱۵ دقیقه است.
- نوشتن نام و نام خانوادگی و شماره دانشجویی بر روی برگه پاسخنامه فراموش نشود.
- آزمون کتاب و جزوه بسته است و تنها استفاده از دو برگه A4 شامل روابط ریاضی مجاز است.
- مسیر آپلود پاسخنامه CW است. در صورت اضطرار ولی در راس ساعت مقرر می‌توانید به آدرس ectexam@gmail.com ارسال نمایید.
- پاسخنامه را به فرمت pdf ارسال کنید.
- نام پاسخنامه کامل به شکل 99xxxxxx.pdf باشد. (لطفا پاسخنامه به ترتیب سوالات تنظیم شود)
- پاسخنامه را بر روی برگه به قطع نزدیک A4 (بدون خط) و با قلمی با رنگ واضح نوشته و عمودی اسکن کنید.
- از ارسال پاسخ نامه‌های متعدد در دو مسیر مشخص شده و ارسال در واتس‌آپ خودداری کنید.
- عدم رعایت هریک از موارد فوق، موجب از میان رفتن، فرصت تقاضای تجدید نظر می‌باشد.

سوال پنجم {۲۰}: شکل روبرو نسخه Unfold یک شبکه بازگشتی را در طی زمان نشان می‌دهد:

X: ورودی

H: متغیر مخفی

O: خروجی واقعی

L: تابع هزینه

T: مقدار مطلوب خروجی

U/V/W: ماتریس‌های وزن از مرتبه مناسب

الف) روابط ریاضی بازگشتی این شبکه را بنویسید.

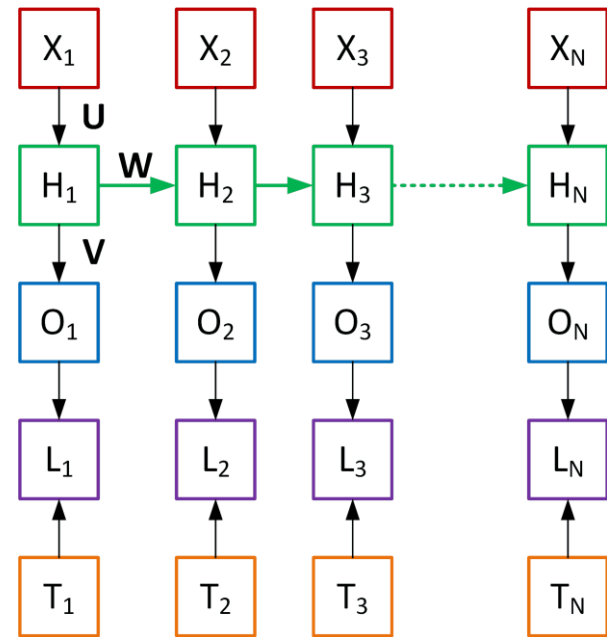
$$H_t = h(WH_{t-1}, UX_t), \quad O_t = g(VH_t)$$

ب) آیا این شبکه دارای مشکل Gradient Vanishing هست؟ به چه دلیل؟ **بله، ارتباط با کل گذشته**

پ) بر اساس روش BPTT، گرادیان لازم برای اصلاح

ماتریس وزن **W** را بنویسید.

$$\sum_{t=2}^{n-1} \left\{ \frac{\partial L_t}{\partial O_t} \frac{\partial O_t}{\partial H_t} \frac{\partial H_t}{\partial W} + \frac{\partial L_{t+1:n}}{\partial H_{t+1}} \frac{\partial H_{t+1}}{\partial H_t} \frac{\partial H_t}{\partial W} \right\}$$



BPTT: BackPropagation Through Time

سوال ششم {۳۵}:

به سوالات زیر پاسخ دهید:

الف) پارامترهای آزاد قابل آموزش شبکه Generator و شبکه Discriminator در فرمول بندی بالا کدام است؟

θ :Generator, w :Discriminator

ب) تابع هدف را به ساده ترین صورت ممکن بر حسب توابع توزیع واقعی: $p(x)$ ، توزیع تولید شده: $q(x)$ ، و تابع Discriminator: $D(x)$ بنویسید.

$$D(x) = \frac{1}{1 + e^{-V_{\omega}(x)}} \Rightarrow e^{-V_{\omega}(x)} = \frac{1 - D(x)}{D(x)}$$

$$F = \int \left\{ p(x) \left(1 - e^{-V_{\omega}(x)} \right) - q(x) \frac{1 - e^{-V_{\omega}(x)}}{1 - \left(1 - e^{-V_{\omega}(x)} \right)} \right\} dx$$

$$F = \int \left\{ p \frac{2D-1}{D} - q \frac{\frac{2D-1}{D}}{1 - \frac{2D-1}{D}} \right\} dx = \int \left\{ p \frac{2D-1}{D} - q \frac{2D-1}{1-D} \right\} dx = \int \left\{ p \left(2 - \frac{1}{D} \right) - q \left(-2 + \frac{1}{1-D} \right) \right\} dx$$

پ) به ازای چه انتخابی از $D(x)$ تابع بند قبل بهینه می شود (در ساده ترین شکل ممکن)

$$\text{Calculus of Variations: } \frac{p}{D^2} - \frac{q}{(1-D)^2} = 0 \Rightarrow \sqrt{p}(1-D) = \pm \sqrt{q}D \Rightarrow D = \frac{\sqrt{p}}{\sqrt{p} \pm \sqrt{q}}$$

$$\Rightarrow D = \frac{\sqrt{p}}{\sqrt{p} + \sqrt{q}}$$

ت) مقدار تابع $F(\theta, \omega)$ را با استفاده از مقدار $D(x)$ حساب شده در بند قبل را در ساده ترین شکل ممکن بنویسید و توجیه کنید.

$$F = \int \left\{ \left(\frac{2D-1}{D} \right) p - \left(\frac{2D-1}{1-D} \right) q \right\} dx = \int \left(2 \frac{\sqrt{p}}{\sqrt{p} + \sqrt{q}} - 1 \right) \left(\frac{p}{\frac{\sqrt{p}}{\sqrt{p} + \sqrt{q}}} - \frac{q}{\frac{\sqrt{q}}{\sqrt{p} + \sqrt{q}}} \right) dx$$

$$F = \int \left(\frac{\sqrt{p} - \sqrt{q}}{\sqrt{p} + \sqrt{q}} \right) \left((\sqrt{p} + \sqrt{q})\sqrt{p} - (\sqrt{p} + \sqrt{q})\sqrt{q} \right) dx$$

$$F = \int (\sqrt{p} - \sqrt{q})^2 dx$$

در واقع تابع هدف در حالت بهینه طراحی تمایزگر، فاصله دو توزیع را نشان می دهد که مولد باید این فاصله را صفر کند.

ث) در صورت آموزش دقیق و بهینه شبکه Generator مقدار $D(x)$ چه میزان خواهد شد.

$$D = \frac{\sqrt{p}}{\sqrt{p} + \sqrt{q}} = 0.5$$