به‌ نام خدا



|  |  |  |
| --- | --- | --- |
|  |  | یادگیري ژرف‌ |
|  |  | نیم‌ سال اول ٠٣- ١٤٠٢ |
| دانشکدهي مهندسی‌ کامپیوتر |  | مدرس: دکتر حمید بیگی‌ |
|  |  |  |
| تمرین‌ سري اول (١٠٠ نمره) | عنوان تمرین‌ | زمان تحویل‌: ١١ آذر |

لطفا نکات زیر را رعایت‌ کنید:

* سوالات خود را از طریق‌ پست‌ مربوط به‌ تمرین‌ در Quera مطرح کنید.
* در هر کدام از سوالات، اگر از منابع‌ خارجی‌ استفاده کردهاید باید آن را ذکر کنید. در صورت همفکري با افراد دیگر هم‌ باید نام ایشان را در سوال مورد نظر ذکر نمایید.
* پاسخ‌ ارسالی‌ واضح‌ و خوانا باشد. در غیر این‌ صورت ممکن‌ است‌ منجر به‌ از دست‌ دادن نمره شود.
* پاسخ‌ ارسالی‌ باید توسط‌ خود شما نوشته‌ شده باشد. به‌ اسکرین‌ شات از منابع‌ یا پاسخ‌ افراد دیگر نمرهاي تعلق‌ نمی‌گیرد.
* در صورتی‌ که‌ بخشی‌ از سوال ها را جاي دیگري آپلود کرده و لینک‌ آن را قرار داده باشید، حتما باید تاریخ‌ آپلود مشخص‌ و قابل‌ اعتنا باشد.

ـ تمام پاسخ‌ هاي خود را در یک‌ فایل‌ با فرمت‌ HW [SID] [Fullname].zip روي کوئرا قرار دهید.

سوال ١: (نظري) عنوان سوال اول (٣٠ نمره)

١) با توجه‌ به‌ معماري LSTM داده شده در شکل‌ ١ به‌ سوالات پاسخ‌ دهید.

(آ) معادله‌ هاي تمام متغیر هاي خروجی‌ این‌ سلول را بنویسید.

(ب) *@hT* را حساب کنید.

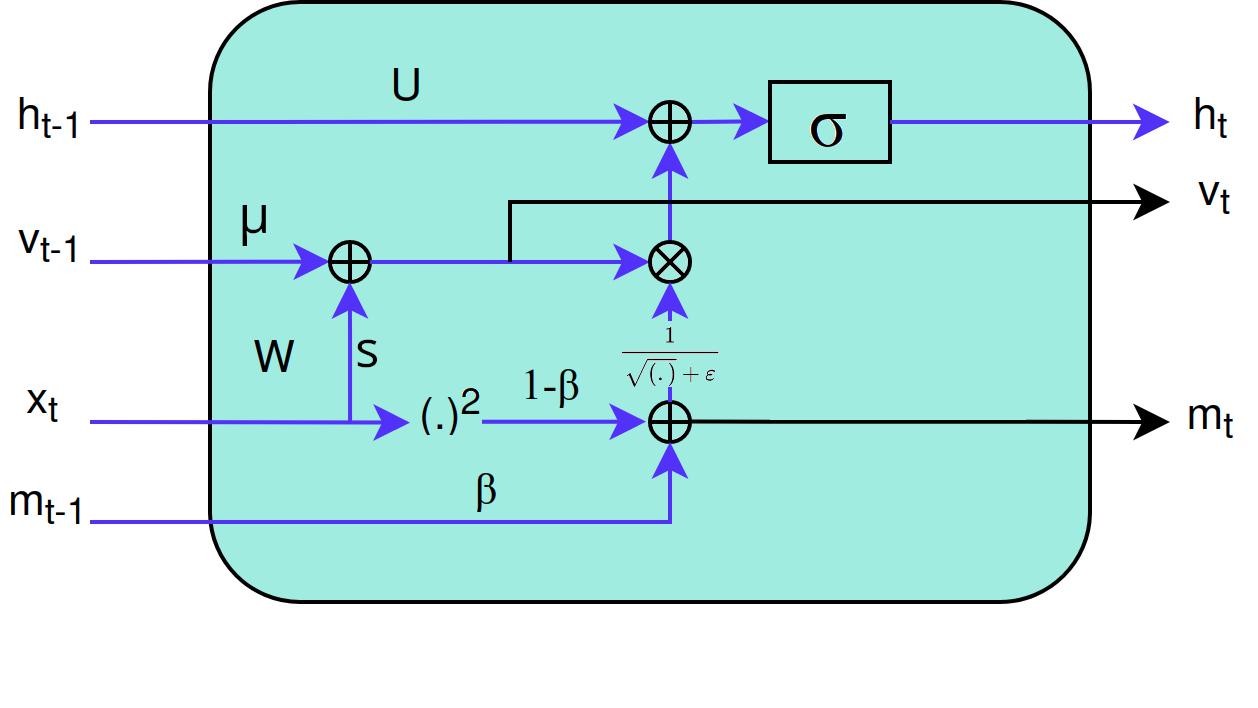
*xi*

(ج) *@hT* را هم‌ حساب کنید

*@hi*

(د) این‌ معماري با مشکل‌ Gradient Vanishing یا Gradient Explosion رو به‌ رو می‌ شود؟ توضیح‌ دهید و راه حلی‌ براي آن ارائه‌ دهید.

(ه) این‌ سلول شبیه‌ به‌ چه‌ الگوریتم‌ بهینه‌سازي عمل‌ می‌کند؟ نحوه کار این‌ الگوریتم‌ را توضیح‌ دهید.



شکل‌ ١: LSTM

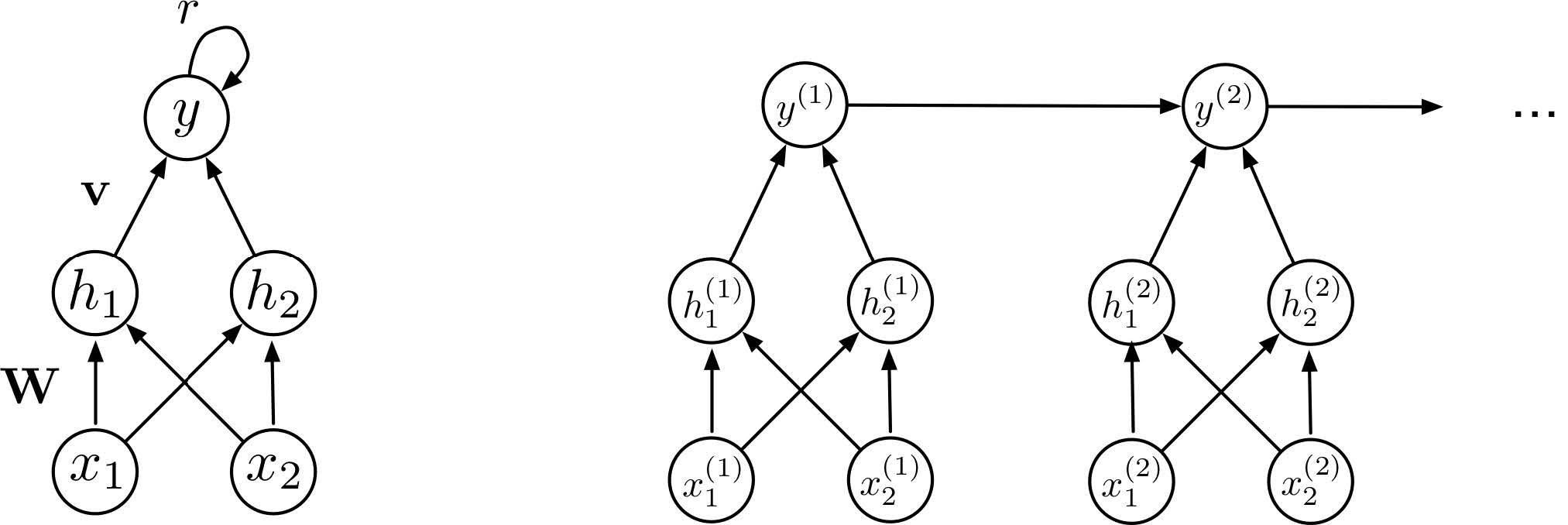
سوال ٢: (نظري) عنوان سوال دوم (٢٠ نمره)

در شکل‌ زیر معماري یک‌ شبکه‌ بازگشتی‌ (راست‌) و یک‌ سلول از آن (چپ‌) را مشاهده می‌کنید. می‌ خواهیم‌ این‌ RNN را به‌ نحوي طراحی‌ کنیم‌ که‌

( ) ( )

با دریافت‌ دو دنباله‌ دودویی‌ ورودي ( *N*١*. . . , x* ( ,(١)١x1 = *x* و ( *N*٢*. . . , x* ( ,(١)٢x2 = *x* مشخص‌ کند که‌ آیا آن ها یکسان هستند یا نه‌.

١



در این‌ شبکه‌، محاسباتی‌ که‌ در هر گام زمانی‌ انجام می‌ شود به‌ صورت زیر است‌:

*h*(*t*)= *φ*(*W x*(*t*)+ *b*)

|  |  |  |
| --- | --- | --- |
| ١ + *c*) *t >* (١ |  | *φ*(*vT h*(*t*)+ *ry*(*t* |
| ١ *t* = | *−* | (٠*y*(*t*) = \_*φ*(*vT* *h*(*t*) + *c* |

که‌ در آن *φ* تابع‌ فعالسازي پله‌، *W* ماتریس‌ 2 *×* 2، *b* و *v* بردار 2 بعدي، *r*، *c* و ٠*c* اسکالر می‌باشند.

در این‌ شبکه‌ در هر گام زمانی‌ *t*، عناصر *t*)١*x*( و *t*)٢*x*( به‌ عنوان ورودي به‌ مدل داده می‌ شوند و خروجی‌ *y*(*t*) محاسبه‌ می‌کند که‌ آیا همه‌ جفت‌ عناصر تا زمان *t* مطابقت‌ داشته‌اند یا خیر. به‌ این‌ ترتیب‌ خروجی‌ ( *y*(*N* نشان دهنده برابري / نابرابري دو دنباله‌ دودویی‌ ورودي خواهد بود.

پارامترهاي *W*، *b*، *v*، *r*، *c* و ٠*c* را به‌ نحوي تعیین‌ کنید که‌ شبکه‌ داده شده تابع‌ مورد نظر را پیادهسازي کند.

سوال ٣: (عملی‌) تکمیل‌ نوت بوك ها (٥٠ نمره)

توضیحات لازم در مورد هر نوت بوك داخل‌ آن وجود دارد. لطفا آن ها را تکمیل‌ کنید.

٢