

به نام خدا

## تمرین سری سوم درس رباتیک اجتماعی و شناختی



(شبکه های عصبی بازگشتی و مدل های مخفی مارکوف)

مدرس: علیرضا طاهری

دانشکده مهندسی مکانیک، دانشگاه صنعتی شریف

۱- یک سیستم دینامیکی (به عنوان مثال یک سیستم مکانیکی) دلخواه را تحت ورودی های متغیر با زمان در نظر بگیرید. ابتدا آن را به صورت تحلیلی و یا با بهره گیری از روش های عددی (نظیر ode45 در متلب) حل کرده، سپس پاسخ های به دست آمده را به بخش های دادگان آموزش و دادگان آزمون تقسیم کنید. یک شبکه بازگشتی مناسب آموزش دهید تا این سیستم دینامیکی را مدل کند. ورودی های شبکه را در هر گام زمانی، ورودی سیستم دینامیکی (به عنوان مثال نیروها/گشتاورهای وارد به سیستم در هر لحظه، موقعیت اعمال آن نیروها و گشتاورها، و در صورت تمایل و به صورت عمومی تر، جرم ها و ممان های اینرسی اجزا و ...) و همچنین خروجی لحظه قبل (به عنوان مثال موقعیت ها و سرعت های خطی/زاویه ای گام پیش) در نظر بگیرید.

الف) خروجی شبکه را از  $t=0$  تا زمان دلخواه رسم کرده و دقت شبکه خود را گزارش نمایید. قدرت شبکه خود را در دادگانی در خارج از مجموعه آموزش خود (به عبارتی در زمان هایی بیش از حداکثر زمان آموزش شبکه) نیز بررسی کنید؛ پیش تر مشاهده کردیم که شبکه های عصبی چندلایه معمولی، لزوماً برونمایی مناسبی ندارند. آیا با بهره گیری از شبکه های عصبی بازگشتی، این شرایط ارتقا یافته است؟

ب) با تغییر ورودی سیستم دینامیکی و بدون آموزش مجدد شبکه، بررسی کنید که آیا شبکه همچنان قادر به یافتن خروجی های صحیح می باشد یا خیر.

یک پیشنهاد برای سیستم دینامیکی، یافتن موقعیت و سرعت یک جسم متصل به یک فنر و دمپر تحت نیروی نوسانی غیرخطی (پالس مربعی) بر حسب زمان است:

$$m\ddot{x} + c\dot{x} + kx = f(t)$$

$$f(t) = \begin{cases} 1 & 0 \leq t < 1 \\ 0 & 1 \leq t < 2 \end{cases}, \quad f(t+2) = f(t)$$

اگرچه شما می توانید به دلخواه سیستم های دینامیکی جذاب تری (!) را نیز مدل کنید.

۲- با جستجوی اینترنتی و به صورت کیفی به پرسش های زیر پاسخ دهید:

الف) در مدل های مارکوف، تاکنون فرض ما بر وابستگی حالت ها (states) فقط به یک گام زمانی قبل بود. فرض کنید باریدن باران امروز، بستگی به شرایط هر دو روز گذشته خود دارد. ماتریس انتقال حالت (State Transition Probability) در این وضعیت چگونه است؟

ب) در شبکه های بازگشتی دنباله-به-دنباله (sequence-to-sequence)، از دو روند متفاوت برای آموزش و تست استفاده می شود. در فاز آموزش، در مرحله  $t$  از برچسب واقعی مرحله  $t-1$  استفاده می شود و شبکه، خروجی درست مرحله قبل را استفاده می کند. در فاز تست، از خروجی مدل در مرحله  $t-1$  به عنوان ورودی مرحله  $t$  استفاده می شود. این تفاوت باعث وجود مشکل Exposure Bias می شود. این مشکل را توضیح دهید. چه راه حلی برای آن ارائه شده است؟

پ) یکی از مشکلات شبکه های عصبی عمیق، می تواند بیش برآزش (فیت شدن بیش از حد آن ها بر روی داده های آموزشی) باشد. از این رو برای آموزش آن ها، می توان از تکنیک های منظم ساز مثل Batch Normalization استفاده کرد که در تمرین سری دوم با آن آشنا شدید. بیان کنید که آیا می توان به طور مستقیم از این روش در شبکه های عصبی بازگشتی نیز استفاده کرد؟ چه مشکلاتی برای استفاده از آن وجود دارد و راه حل چیست؟

۳- مجموعه مدل های مخفی مارکوف مناسب برای شش حرف فارسی «ب»، «ت»، «ح»، «ص»، «ع» و «ک» ارائه کنید (هر حرف یک HMM). جهت جمع آوری دادگان می توانید از کدهای ارائه شده ما در سامانه درس افزار بهره گیرید. برای هر حرف حداقل از ۵ نمونه آموزشی استفاده کنید. جهت آموزش مدل مخفی مارکوف خود نیز می توانید از کدهای متلب ارائه شده توسط دکتر کلینون استفاده نمایید. الف) برای هر مدل،  $A$ ،  $B$  و  $\pi$  را گزارش نمایید. ب) دقت مدل های خود را در تعدادی نمونه آزمون (حداقل ۳ نمونه برای هر دسته) با ارائه ماتریس درهم ریختگی ارزیابی نمایید.

۴- لب خوانی خودکار توسط سامانه های رباتیک اجتماعی به صورت بالقوه در تشخیص بهتر علائم زبان اشاره ایرانی کاربرد دارد. (جهت اطلاع شما، متأسفانه به علت عدم آشنایی افراد عادی در جامعه و حتی معلمان افراد ناشنوا/با مشکلات شنوایی در کشور با زبان اشاره ایرانی، افراد ناشنوا از همان ابتدای دوران تحصیل به لب خوانی عادت کرده و (به جای تعامل مستمر دوطرفه با دیگران از طریق حرکات دست و دهان به صورت همزمان)، برای رفع نیازهای خود، ناچار به خیره شدن در زمینه لب خوانی هستند!! بگذریم...) در این سوال مجموعه دادگانی در قالب داده های آموزش و آزمون به شکل فیلم های کوتاه (بدون صدا) با پسوند mp4 در اختیار شما قرار می گیرد که شما می بایست با طراحی شبکه ای مناسب، یک مسئله دسته بندی انجام دهید. تعداد کلمات استفاده شده در این سوال ۵ کلمه (شامل ۱: ایران، ۲: خوشحال، ۳: معلم، ۴: سلام، ۵: خداحافظ) است که توسط ۱۰ فرد مختلف، هر یک به تعداد ۴ مرتبه، بیان شده اند. تعداد افراد آموزش ۸ (شماره های ۱ تا ۸) و تعداد افراد تست ۲ (شماره های ۹ و ۱۰) می باشد. نام گذاری فایل ها بدین صورت است که عدد بعد از S شماره فرد، عدد بعد از W شماره کلمه و عدد بعد از r مرتبه تکرار را نشان می دهد (برای مثال S1-W2-r3 فایل مربوط به تکرار سوم کلمه ۲: خوشحال توسط فرد شماره ۱ است). یک شبکه تلفیقی پیچشی-بازگشتی ارائه دهید تا دسته بندی این ۵ کلمه را انجام دهد. الف) ساختار مورد نظر خود را برای انجام این تکلیف طبقه بندی توضیح دهید. ب) ماتریس درهم ریختگی را بر روی دادگان آزمون ارائه کنید.

لینک دانلود دادگان آموزش و آزمون:

<https://drive.google.com/drive/folders/19po4DgGLAobWcn2Je-hp9lawUWXLYWv2?usp=sharing>

علاوه بر خواسته های مسائل، می بایست تمامی کدهای خود را نیز بارگذاری کنید.