

بسم الله الرحمن الرحيم



دانشگاه صنعتی مالک اشتر

دانشکده مهندسی برق و کامپیوتر

درس یادگیری عمیق

تمرین شماره چهارم : شبکه های عصبی بازگشتی

استاد: آقای دکتر کیوان راد

دستیار آموزشی: آقای سیدعلیان

مهلت تحویل: ۱۴۰۴/۰۳/۵

امیرالمومنین امام علی علیه السلام:

ارزش هر انسانی به آن کاری است که آن را نیکو انجام می دهد.

بهار ۱۴۰۴

تولید موزیک با شبکه عصبی بازگشتی

مقدمه

یکی از کاربردهای جالب شبکه های عصبی بازگشتی^۱ یا RNN ها، استفاده از آنها برای تولید موزیک^۲ است. RNN ها در مدیریت داده های متوالی^۳ خوب عمل می کنند. با توجه به این که در موسیقی هر نوت به نوت قبلی خود وابسته است، می توان از RNN ها برای تولید موزیک استفاده کرد.

قسمت اول: load کردن دیتا

برای این سوال ما از دیتای Classical Music MIDI استفاده خواهیم کرد. فایل دیتایی که در اختیار شما قرار داده می شود دارای آهنگ های آهنگسازان مختلف می باشد. ما فقط از آهنگ های آهنگساز لهستانی [Frédéric Chopin](#) استفاده خواهیم کرد.

برای load کردن دیتا آهنگ های این هنرمند را از پوشه هم نام خودش با استفاده از کتابخانه music21 بخوانید. سپس باید cord ها و note های موجود در هر آهنگ را به دست بیاوریم. برای این کار لازم است که ابتدا یک تعریفی از note و cord در موسیقی به دست بیاورید. در نهایت همه note ها موجود در هر آهنگ را در یک لیست ذخیره کنید. برای extract کردن note ها می توانید از تابع extract_notes داده شده استفاده کنید.

قسمت دوم: پیش پردازش دیتا

در این مرحله پس از این که لیست corpus را به دست آوردید، تعداد تکرار هر note را شمرده و note های کم تکرار^۴ را از corpus حذف کنید. لازم به ذکر است که تعداد حداقل تکرار برای حذف کردن یک note اختیاری است و می توانید این پارامتر را خودتان تعیین کنید. جهت درست کردن دیتای آموزش و تست شبکه می توانید از corpus حاصل از مرحله قبل note ها را به صورت sequence های ۴۰ تایی از corpus خوانده و note ۴۱ امی بعد از sequence را به عنوان لیبل و دیتای تست در نظر بگیرید. برای مثال ایندکس ۰ تا ۳۹ به عنوان یک sequence و ایندکس ۴۰ به عنوان لیبل این sequence در نظر گرفته

¹ Recurrent neural network (RNN)

² Music generation

³ sequential data

⁴ rare

می‌شود. ایندکس ۱ تا ۴۰ به عنوان sequence بعدی و ایندکس ۴۱ به عنوان لیبل این sequence در نظر گرفته میشود و الی آخر ...

پس از ایجاد sequence ها با استفاده از یک دیکشنری sequence ها و لیبل های مورد نظر را encode کنید و در این مرحله دیتای آموزشی و لیبل های شما در دسترس است. توجه کنید که برای آموزش شبکه لیبل ها را باید به صورت one hot encoding تبدیل کنید.

پس از این کار دیتای train و test را جدا کرده و در مرحله بعد شبکه بازگشتی خود را پیاده سازی می‌کنیم.

قسمت سوم: پیاده سازی شبکه عصبی بازگشتی

میتوانید از مدل زیر یا هر مدل دلخواه خود استفاده کنید. از تابع CategoricalCrossentropy یا CrossEntropyLoss برای محاسبه loss استفاده کنید. هم چنین تابع فعال سازی لایه fully connected را softmax در نظر بگیرید. ساختار انتخابی خود را در گزارش شرح دهید.

| Layer (type) | Output Shape | Param # |
|-----------------------------|-----------------|---------|
| lstm (LSTM) | (None, 40, 512) | 1052672 |
| lstm_1 (LSTM) | (None, 256) | 787456 |
| dense (Dense) | (None, 256) | 65792 |
| dense_1 (Dense) | (None, 266) | 68362 |
| ===== | | |
| Total params: 1,974,282 | | |
| Trainable params: 1,974,282 | | |
| Non-trainable params: 0 | | |

قسمت چهارم: ارزیابی شبکه

مقدار loss را برای فرایند آموزش چاپ کنید. یک آهنگ با استفاده از تابع داده شده تولید کرده و آن را به همراه فایل گزارش خود آپلود کنید. آهنگ تولید شده با فرمت midi را میتوانید با استفاده از [ابزارهای آنلای](#) موجود به فرمت wav تبدیل کنید. همین مراحل را برای موزیک هایی که آهنگساز آن [Mozart](#) است انجام دهید. تاثیر dropout بر روی عملکرد شبکه را آزمایش کنید. پس از هر لایه lstm و لایه های fully connected یک لایه dropout اضافه کنید.

ملاحظات ارسال تمرین

- در این تمرین فقط مجاز به استفاده از زبان برنامه نویسی پایتون هستید. اما محدودیتی در استفاده از فریمورک ها و کتابخانه های مختلف آن وجود ندارد.
- انتظار می رود گزارش ساختارمند باشد و طبق اصول صحیح نگارش شود. به عنوان مثال پیشنهاد می شود در گزارش خود برای تصاویر زیرنویس و برای جداول بالانویس اضافه شود و از فهرست مطالب و شماره گذاری صفحات استفاده شود.
- گزارش شما در فرآیند تصحیح از اهمیت ویژه ای برخوردار است. لطفاً تمامی نکات و فرض هایی که برای پیاده سازی ها و محاسبات خود در نظر می گیرید را در گزارش ذکر کنید. الزامی به ارائه توضیح جزئیات کد در گزارش نیست. اما باید نتایج بدست آمده را گزارش و تحلیل کنید.
- دستیار آموزشی ملزم به اجرا کردن کدهای شما نیست. بنابراین هرگونه نتیجه و یا تحلیلی که در شرح سوال از شما خواسته شده است را به طور واضح و کامل در گزارش بیاورید.
- در صورت مشاهده تقلب نمرات تمامی افراد شرکت کننده در آن صفر لحاظ می شود.
- فایل گزارش را در پوشه report و کد ها را در پوشه code قرار داده و سپس آنها را در یک فایل zip در سایت کوئرا بارگذاری نمایید.

با آرزوی موفقیت