به نام خدا



دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

درس شبکه‌های عصبی

استاد صفابخش

تمرین پنجم

علیرضا مازوچی

۴۰۰۱۳۱۰۷۵

سوال ۱

ابتدا نکات کلی در مورد پیاده‌سازی را بیان می‌کنم:

* برای پیاده‌سازی هر دو شبکه به پیاده‌سازی یک سلول SimpleRNN در گیت‌هاب مراجعه شده است و برخی از مقداردهی‌های اولیه و ایده‌های کلی از آن گرفته شده است.
* با توجه به کم بودن داده‌ها تعداد دسته‌ها بر روی 16 تنظیم شده است.
* برای آموزش هر دو شبکه اجازه داده شده است تا حداکثر ۲۵۰ گام آموزش انجام بگیرد و برای جلوگیری از بیش برازش از یک کالبک EarlyStopping‌ استفاده شده است.

نتایج برای شبکه Elman‌ در جدول زیر ذکر شده است:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ردیف | تعداد نورون مخفی | تعداد گام آموزش | دقت مجموعه آموزش | دقت مجموعه اعتبارسنجی | دقت مجموعه ارزیابی |
| ۱ | 8 | 141 | ٪۸۶.۷۳ | ۸۰.۳۴٪ | ۸۹.۲۵٪ |
| ۲ | 32 | 193 | ۹۷.۸۴٪ | ۹۷.۳۸٪ | ۹۵.۷۰٪ |
| ۳ | 64 | 98 | ۹۵.۶۸٪ | ۸۹.۱۳٪ | ۹۰.۳۲٪ |
| ۴ | 128 | 45 | ۹۳.۸۳٪ | ۹۳.۴۸٪ | ۹۰.۳۲٪ |
| ۵ | 512 | 29 | ۸۹.۸۱٪ | ۸۲.۶۱٪ | ۸۸.۱۷٪ |
| 6 | 2048 | 46 | ۹۹.۳۸٪ | ۹۵.۶۵٪ | ۹۶.۷۷٪ |

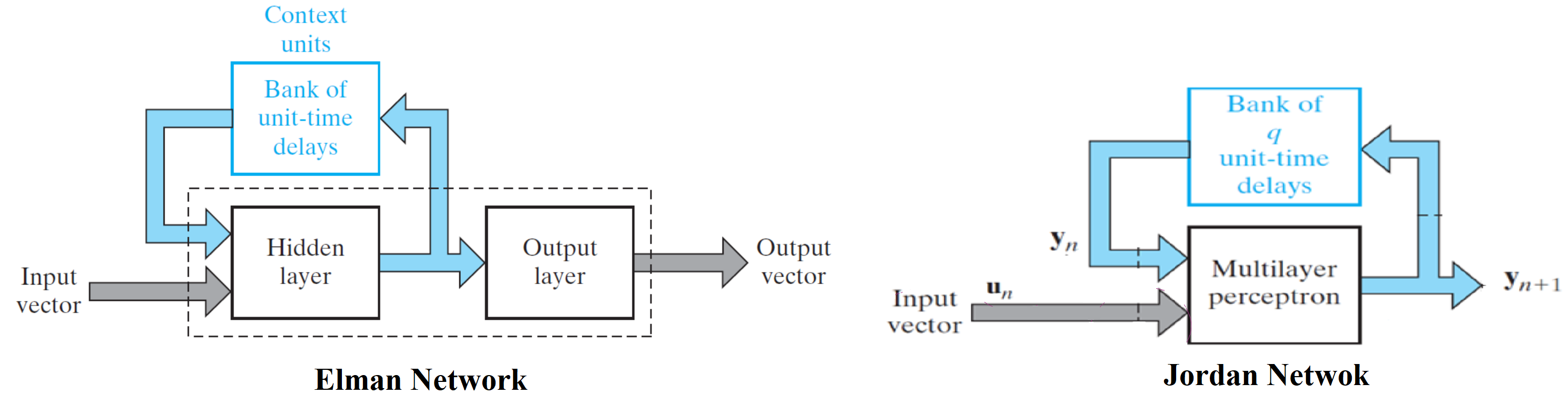
در بین تنظیمات انجام‌شده بهترین دقت ارزیابی را شبکه با 2048 نورون داشته است. البته شبکه با 32 نورون هم توانسته است دقت نزدیکی را ثبت کند. به طور کلی اگر تعداد نورون‌ها خیلی کم باشد ( مثلا 8 نورون) شبکه دچار کم‌برازش می‌شود و نمی‌تواند دقت‌های خوبی را داشته باشد. از طرفی اگر تعداد نورون‌ها زیاد باشد استعداد بیش‌برازش وجود دارد ولی باتوجه به کالبک استفاده‌شده آموزش خاتمه پیدا می‌کند؛ بدین ترتیب جلوی بیش‌برازش تا حد زیادی گرفته می‌شود و به طور جانبی آموزش هم زودتر تمام می‌شود. پس تعداد گام آموزش برای شبکه با تعداد سلول بیشتر، کمتر است.

برای شبکه Jordan به طور مشابه نتایج در جدول زیر آورده شده است:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ردیف | تعداد نورون مخفی | تعداد گام آموزش | دقت مجموعه آموزش | دقت مجموعه اعتبارسنجی | دقت مجموعه ارزیابی |
| ۱ | 8 | 33 | ٪۷۵.۰۰ | ۷۶.۰۹٪ | ۸۱.۷۲٪ |
| ۲ | 32 | 157 | ۹۵.۳۷٪ | ۹۱.۳۰٪ | ۹۴.۶۲٪ |
| ۳ | 64 | 74 | ۹۱.۰۵٪ | ۷۳.۹۱٪ | ۹۱.۴۰٪ |
| ۴ | 128 | 118 | ۹۷.۲۲٪ | ۸۶.۹۶٪ | ۸۹.۲۵٪ |
| ۵ | 512 | 27 | ۸۴.۸۸٪ | ۶۵.۲۲٪ | ۸۶.۰۲٪ |
| 6 | 2048 | 16 | ۷۹.۹۴٪ | ۷۶.۰۹٪ | ۸۱.۷۲٪ |

برای این شبکه هم تعداد نورون بسیار کم مطابق انتظار خوب نیست و شبکه نمی‌تواند آموزش ببیند؛ در کنار آن می‌بینیم که تعداد نورون زیاد هم باعث عدم یادگیری و خاتمه زودهنگام آموزش می‌شود. بهترین حالت برای این شبکه با ۳۲ نورون رخ می‌دهد.

به طور کلی شبکه Elman‌ دقت‌های بهتری از شبکه Jordan دارد. امام تعداد گام و زمان اجرای Jordan مناسب‌تر است. برای فهمیدن چرایی این موضوع باید تفاوت این دو شبکه را درک کنیم. در تصویر زیر معماری دو شبکه آورده شده است:



همانطور که مشخص است در Elman‌ خروجی لایه مخفی وارد شبکه در لحظه بعد می‌شود، اما در Jordan خروجی نهایی شبکه وارد سلول در لحظه بعد می‌شود؛ حال توجه کنید که مسئله ما یک مسئله دو کلاسه است پس خروجی دو بعدی است درحالی که خروجی لایه مخفی به تعداد نورون‌ها ابعاد دارد؛ یعنی اگر 32 نورون داشته باشیم، خروجی ۳۲ بعدی وارد شبکه در لحظه بعد می‌شود. با این اوصاف طبیعی است که Elman‌ می‌تواند دانش بیشتری را از لحظات قبل در اختیار بگیرد و منطقی است که دقت بیشتری داشته باشد و در ازای آن پیچیدگی بیشتر و زمان اجرای بیشتر داشته باشد. پیچیدگی کمتر Jordan در عین حال که باعث دقت کمتر می‌شود، برای تنظیمات با تعداد نورون مخفی بیشتر به طرز شدیدتری منجر به بیش‌برازش می‌شود.

با توجه نتایج دریافت‌شده و تحلیل ارائه‌شده واضحا شبکه Elman برای مجموع‌داده سوال مناسب‌تر به نظر می‌آید.