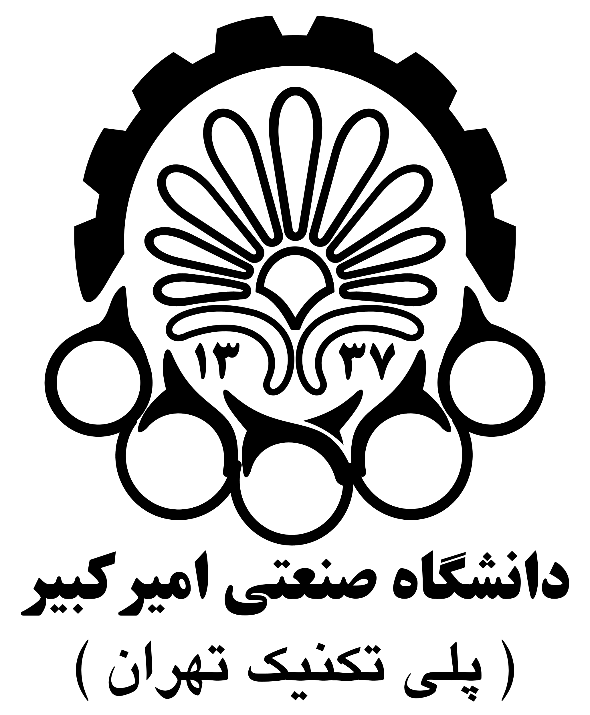
****

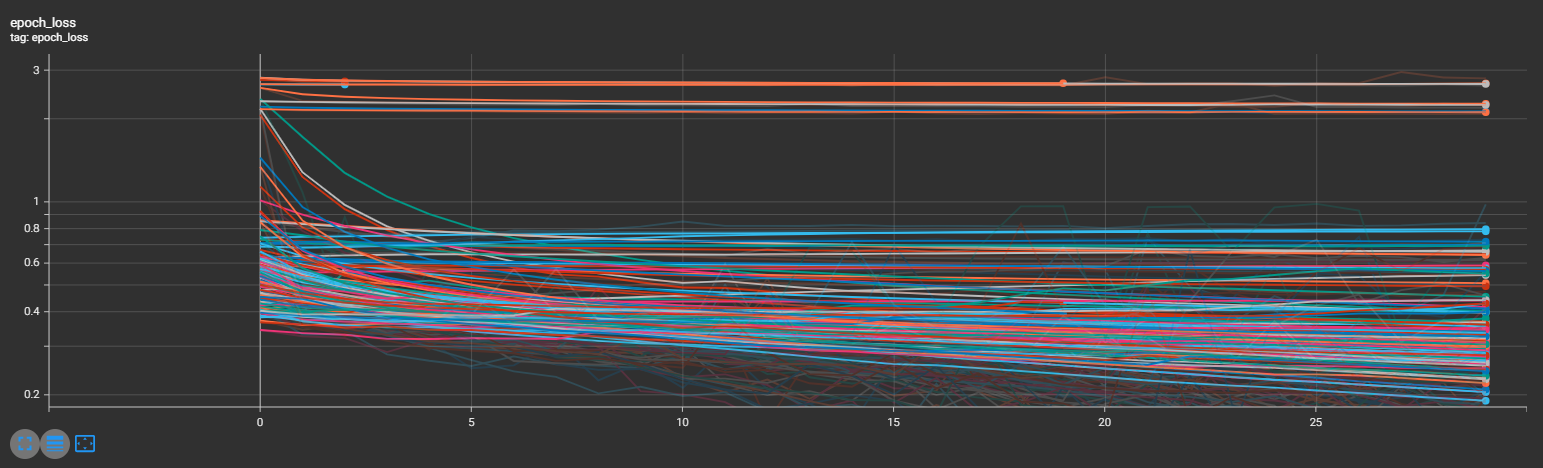
**دانشکده مهندسی کامپیوتر**

**تمرین پنجم درس شبکه عصبی**

**دکتر صفابخش**

**غلامرضا دار 400131018**

**بهار 1401**



فهرست مطالب

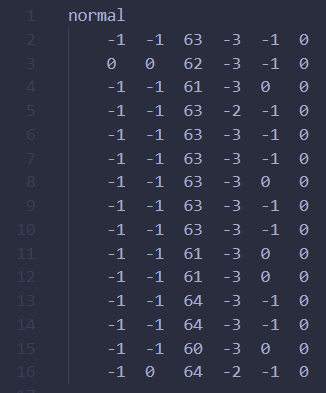
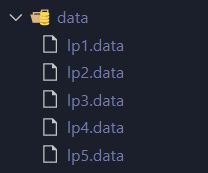
[مقدمه) 3](#_Toc104490423)

[سوال 1) 4](#_Toc104490424)

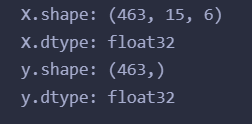
[سوال 2) 7](#_Toc104490425)

# مقدمه)

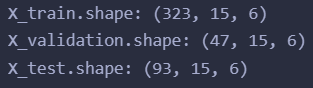
**آماده سازی داده:** قبل از شروع پیاده سازی شبکه های المن و جردن لازم است کمی با دیتاست داده شده در این سوال آشنا شویم. همانطور که در تصویر زیر مشاهده می کنید اطلاعات ذخیره شده توسط سنسور ها به صورت دسته های ۱۵ تایی جدا شده اند و به هر دسته ۱۵ تایی از داده ها یک Label نسبت داده شده است.



همچنین دیتا در ۵ فایل مجزا ذخیره شده است این پنج فایل را خط به خط می­خوانیم و خطوط عددی و لیبل را جدا می کنیم در نهایت به یک دیتاست به ابعاد زیر می رسیم که در مراحل بعد برای آموزش و ارزیابی مدل ها استفاده می شود.



در مرحله بعد طبق خواسته سوال داده ها را به مجموعه‌های Train, Test, Validation تقسیم می­کنیم.



**پیاده سازی شبکه ها:** در ادامه به پیاده­سازی شبکه های جردن و المن با بهره­گیری از کلاس پدر Model از کتابخانه Keras می پردازیم. با مراجعه به نوت بوک تمرین میتوانید جزئیات پیاده سازی این دو شبکه را مشاهده کنید.

# سوال 1)

پس از انجام پیاده­سازی های لازم آزمایش هایی برای پیدا کردن تعداد لایه مخفی بهینه و همچنین مقایسه دو مدل انجام می­دهیم. نتایج آزمایش ها را می‌توانید در جدول زیر مشاهده کنید.

|  |  |  |  |
| --- | --- | --- | --- |
| نتایج آزمایش­های شبکه **المن** | | | |
| صحت Test | صحت Validation | صحت Train | توضیحات مدل |
| 0.8602 | 0.9362 | 0.8947 | المن با 10 نورون مخفی |
| 0.8817 | 0.9362 | 0.9226 | المن با 25 نورون مخفی |
| 0.8924 | 0.9362 | 0.9536 | المن با 50 نورون مخفی |
| **0.9462** | **0.8723** | **0.9783** | **المن با 75 نورون مخفی** |
| 0.9032 | 0.9362 | 0.9783 | المن با 100 نورون مخفی |
| 0.8924 | 0.8723 | 0.9133 | المن با 200 نورون مخفی |

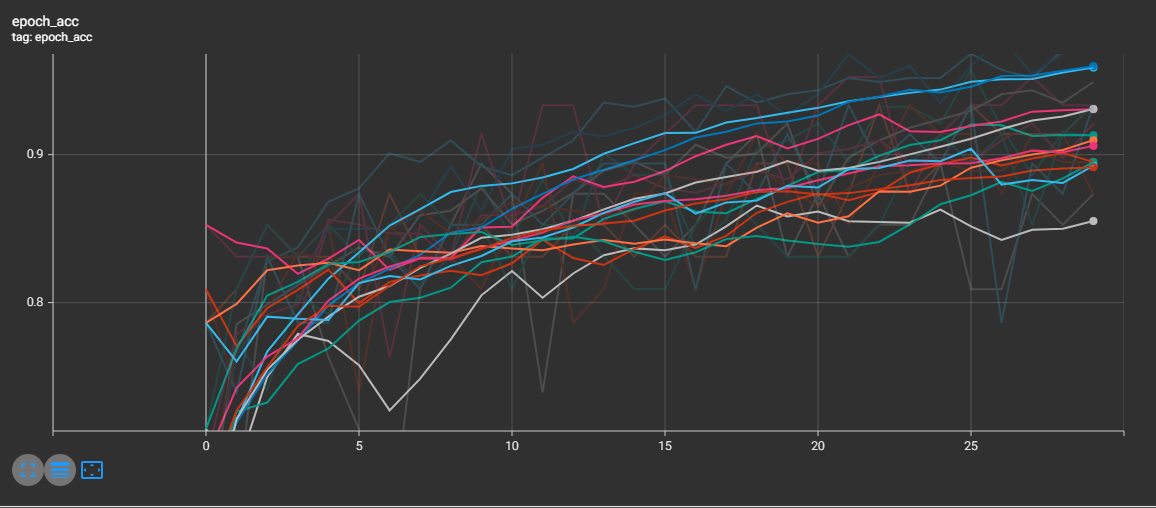
همانطور که مشاهده میکنید افزایش تعداد لایه های مخفی در شبکه المن تا حدی باعث افزایش صحت مدل میشود اما از یک حدی به بعد افزایش تعداد لایه های مخفی صحت مدل را کاهش میدهد.

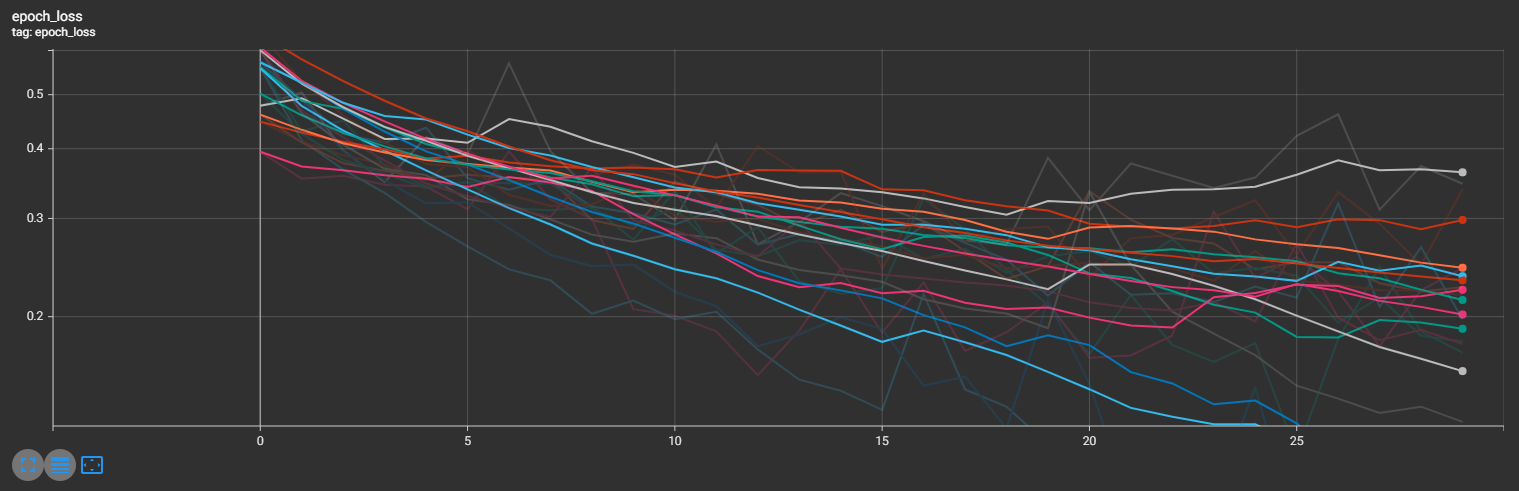
|  |  |  |  |
| --- | --- | --- | --- |
| نتایج آزمایش­های شبکه **جردن** | | | |
| صحت Test | صحت Validation | صحت Train | توضیحات مدل |
| 0.8602 | 0.8723 | 0.8916 | جردن با 10 نورون مخفی |
| 0.8817 | 0.8936 | 0.8793 | جردن با 25 نورون مخفی |
| 0.8924 | 0.8298 | 0.9257 | جردن با 50 نورون مخفی |
| **0.8924** | **0.9149** | **0.9412** | **جردن با 75 نورون مخفی** |
| 0.8064 | 0.8936 | 0.8576 | جردن با 100 نورون مخفی |

مدل جردن نیز مانند مدل المن با افزایش تعداد نورون مخفی تا حدی بهبود می­یابد اما از حدی به بعد دچار کاهش میزان صحت می­شود. تعداد نورون بهینه برای هر دوی این شبکه ها عدد **75** بود.

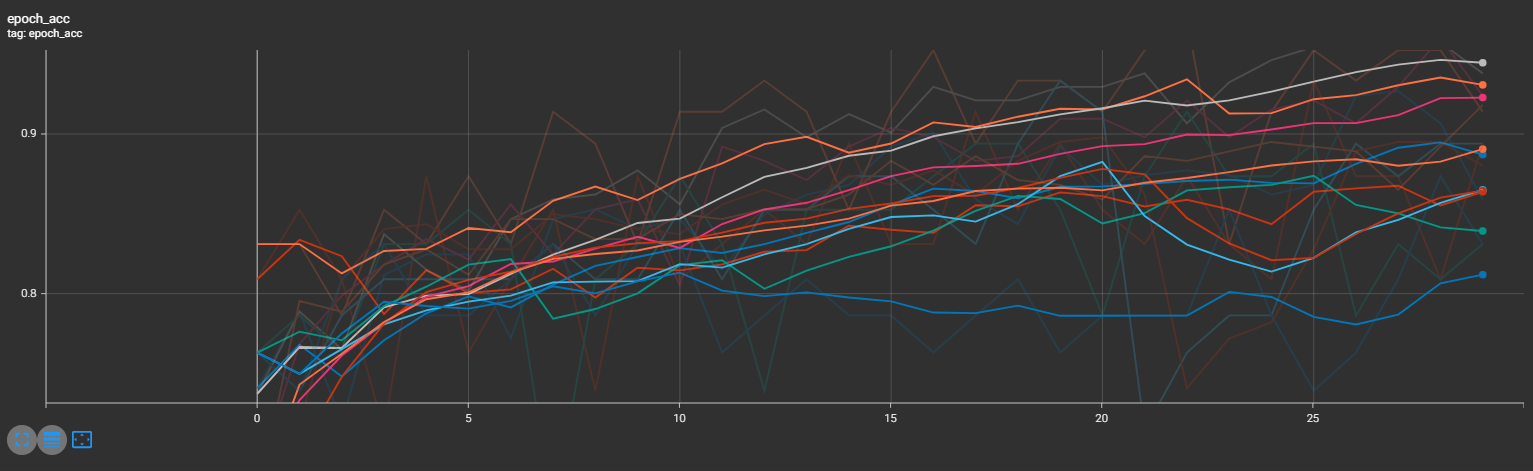
در ادامه نمودار Accuracy و Loss مربوط به تعدادی از آزمایش های مربوط به شبکه های المن و جردن را مشاهده میکنید. مشاهده می­شود که مدل المن همگرایی بهتری دارد و مدل جردن در بعضی حالات رفتارهای ناهنجاری دارد.

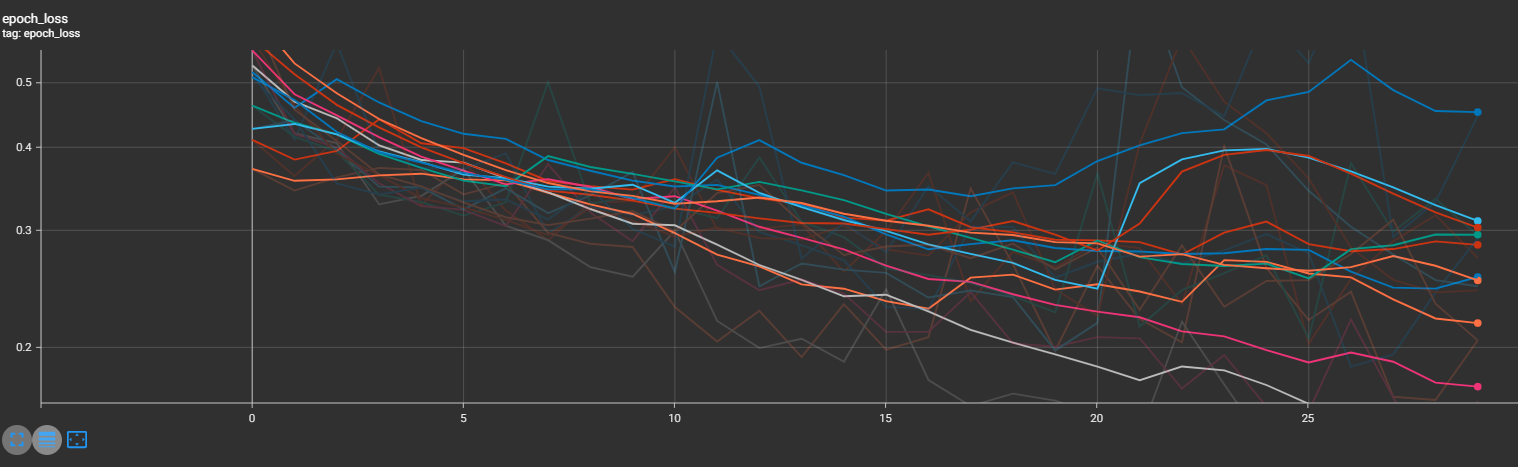
**مدل المن**

****

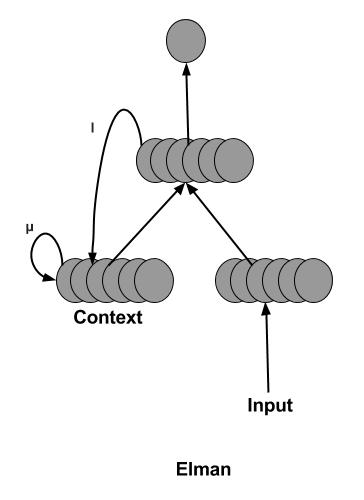
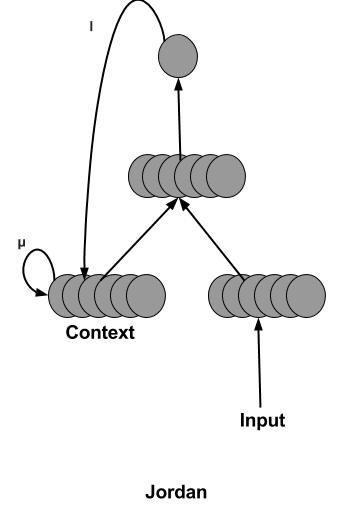
****

**مدل جردن**

****

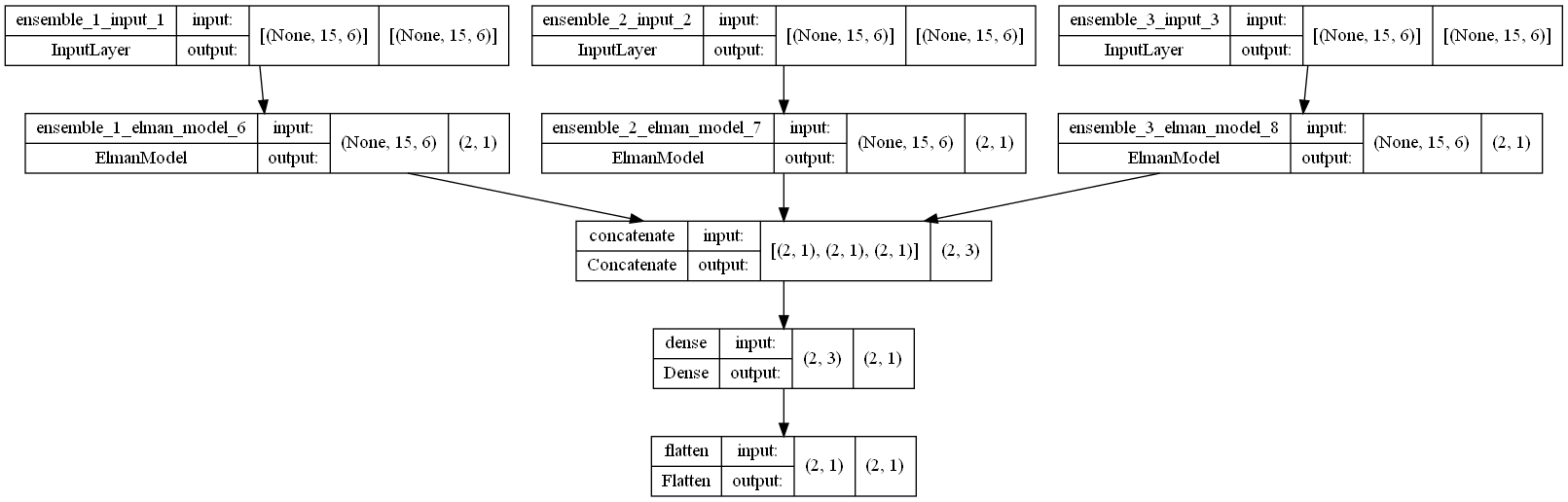
****

**مقایسه نتایج المن و جردن:** همانطور که در تصاویر زیر میبینید شبکه علمن خروجی لایه مخفی مرحله قبل را به ورودی لایه مخفی مرحله بعد می‌دهد .ما شبکه جردن خروجی مدل را برمیگرداند و به ورودی لایه مخفی مرحله بعد می­دهد. با توجه به نتایج به دست آمده مدل **المن** برای این مسئله **مناسب­تر** است. اگر از دید تاریخی نیز به قضیه بنگریم، مدل المن به عنوان مدل جایگزینی برای مدل جردن معرفی شده بود.



# سوال 2)

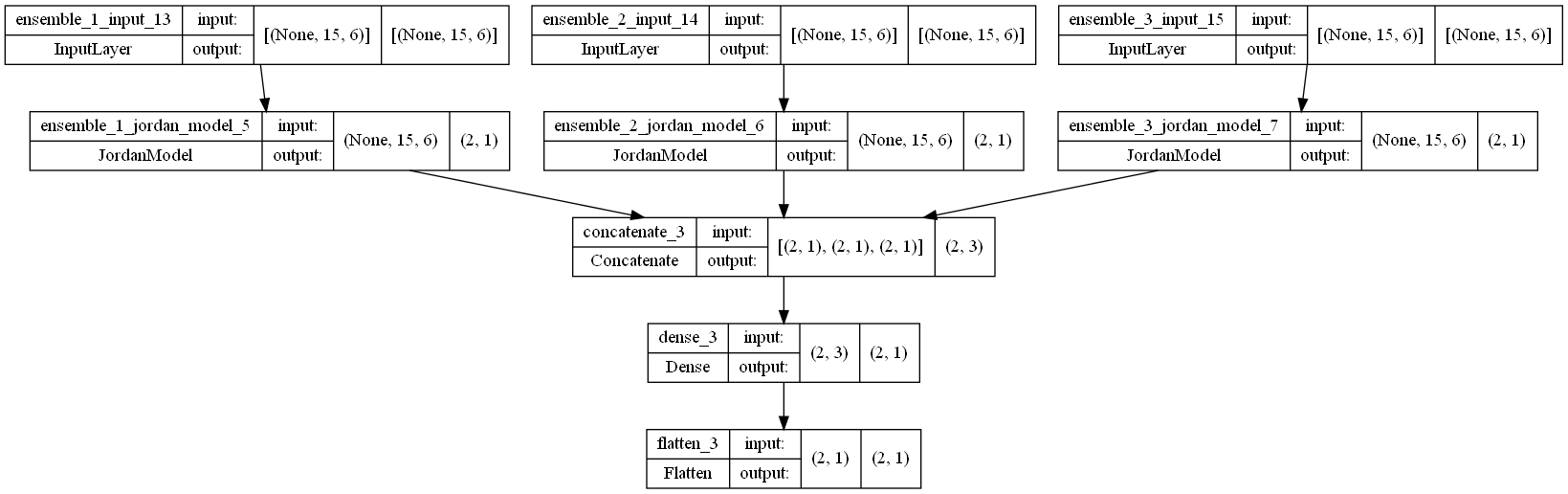
بنابر خواسته سوال ابتدا مدل المن را به صورت گروهی در می­آوریم. مدل گروهی حاصل را در تصویر زیر مشاهده می کنید.



این مدل از ۳ زیرمدل المن تشکیل شده است این زیرمدل ها ورودی یکسانی می­گیرند و در نهایت مدل گروهی تصمیم می­گیرد که این مدل ها چگونه با هم ترکیب شوند تا به بهترین نتیجه دست یابد (لزوما میانگین یا مُد گرفته نمی‌شود). نتیجه تعدادی از آزمایش های این مدل گروهی را در جدول زیر مشاهده می کنید.

|  |  |
| --- | --- |
| نتایج آزمایش­های شبکه گروهی با زیرمدل های **المن** | |
| صحت Test | تعداد نورون­های مخفی زیرمدل­ها |
| **0.9139** | **[10, 25, 50]** |
| 0.8817 | [10, 25, 50, 75] |
| 0.9032 | [10, 25, 50, 75, 100] |

پس از شبکه المن نوبت به مدل گروهی ساخته شده توسط شبکه های جردن رسیده است. مانند مدل المن این مدل گروهی را با زیرمدل های جردن می­سازیم. تصویر زیر گراف یک مدل با 3 زیر مدل جردن است.

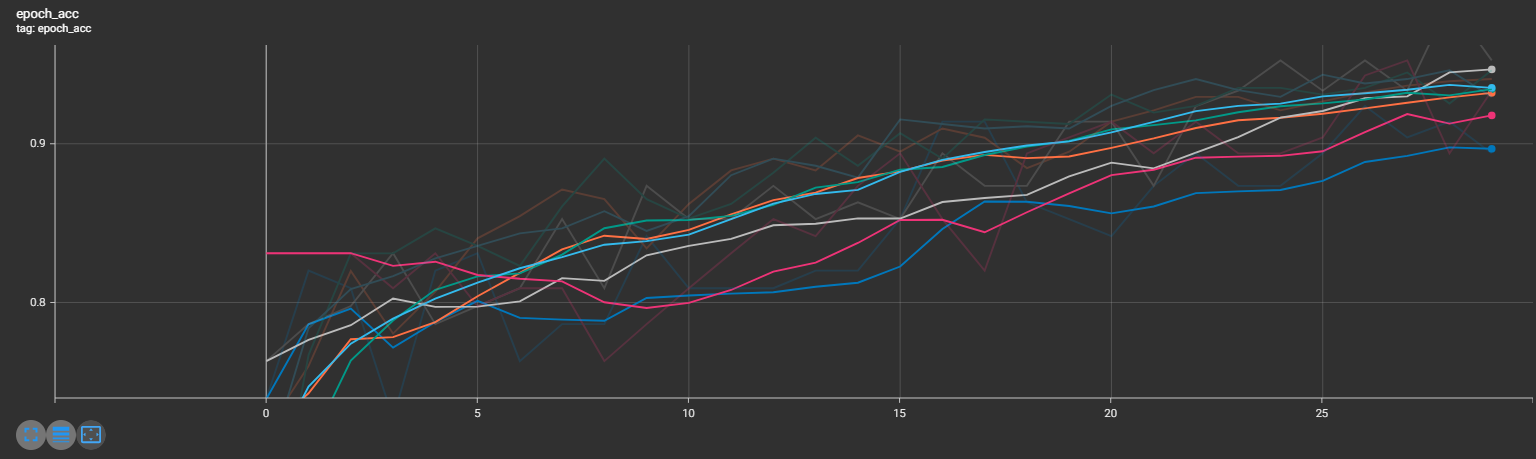


نتیجه تعدادی از آزمایش های مدل گروهی ساخته شده با زیرمدل های جردن را در جدول زیر مشاهده میکنید.

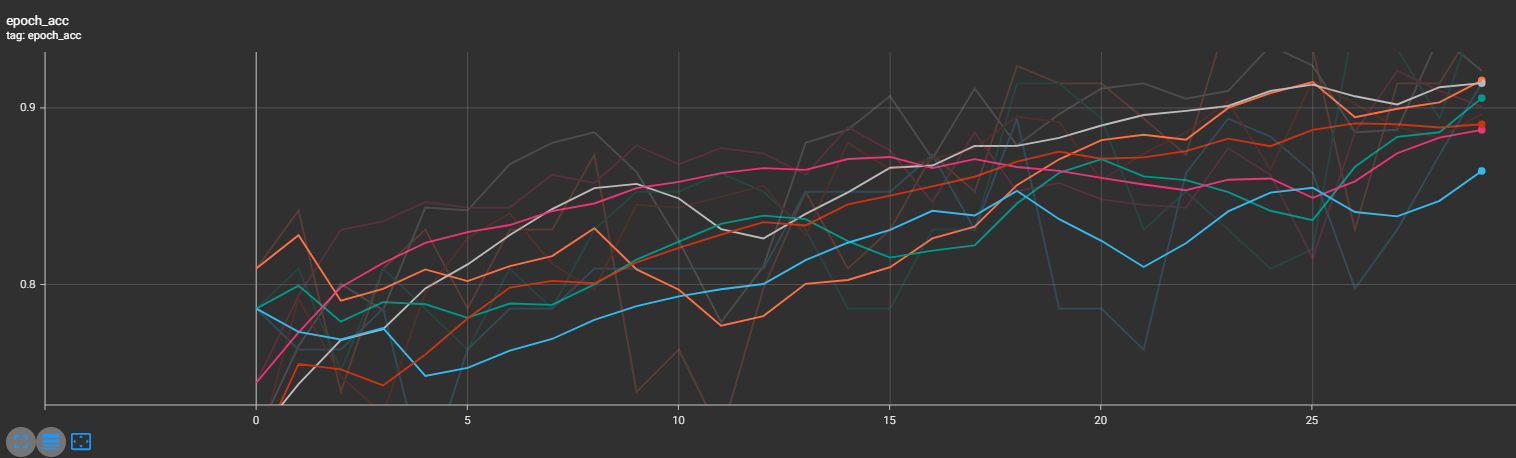
|  |  |
| --- | --- |
| نتایج آزمایش­های شبکه گروهی با زیرمدل های **جردن** | |
| صحت Test | تعداد نورون­های مخفی زیرمدل­ها |
| 0.8709 | [10, 25, 50] |
| **0.9139** | **[10, 25, 50, 75]** |
| 0.8602 | [10, 25, 50, 75, 100] |

در این جا نیز پترن مشابهی مشاهده میکنیم. افزایش تعداد زیرمدل ها تا حدی به بهبود صحت مدل کمک میکنند اما از یک حدی به بعد باعث کاهش آن میشوند. بهترین مدل های گروهی برای هر دو حالت با **رنگ سبز** مشخص شده اند.

**نمودار Accuracy مدل های گروهی با زیرمدل های المن**

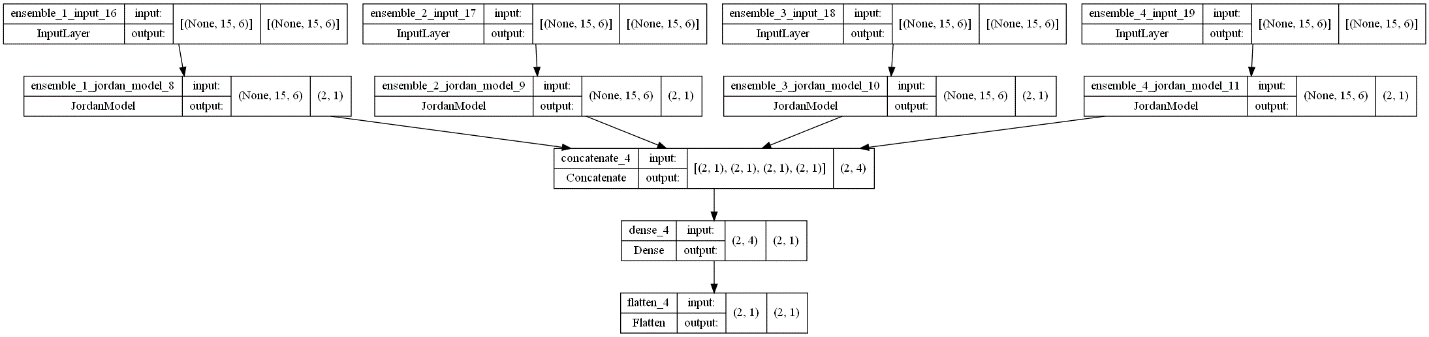


**نمودار Accuracy مدل های گروهی با زیرمدل های جردن**

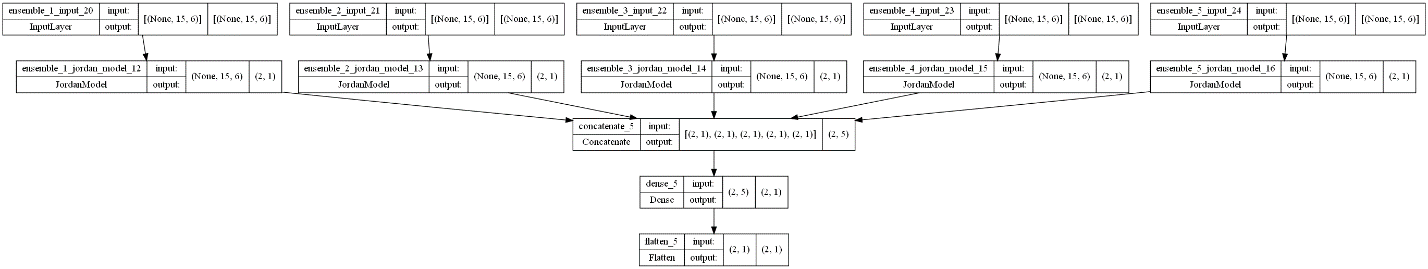


تصاویر مربوط به گراف سایر مدل های گروهی در فایل ارسالی تمرین موجود است. چهار نمونه دیگر از این گراف ها را مشاهده میکنید.

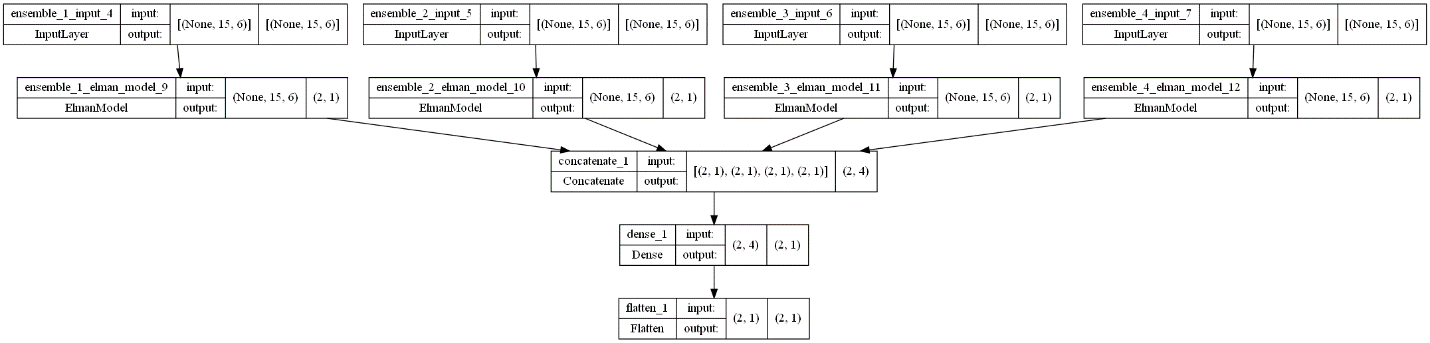
جردن با 4 زیرمدل



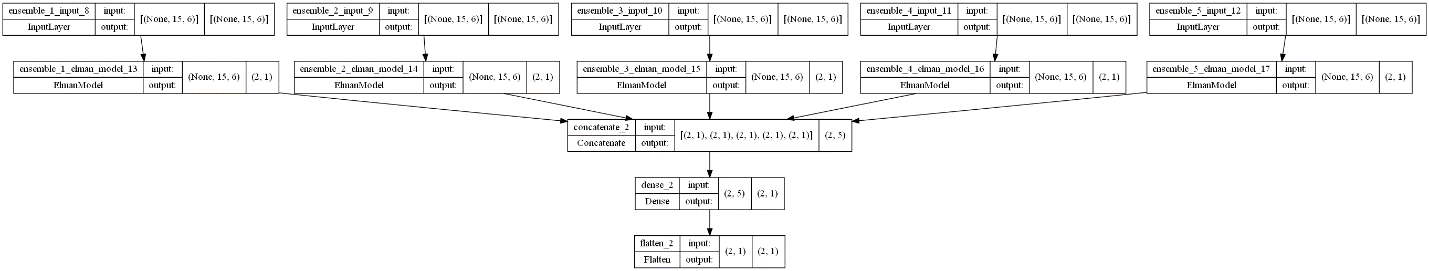
جردن با 5 زیرمدل



المن با 4 زیرمدل

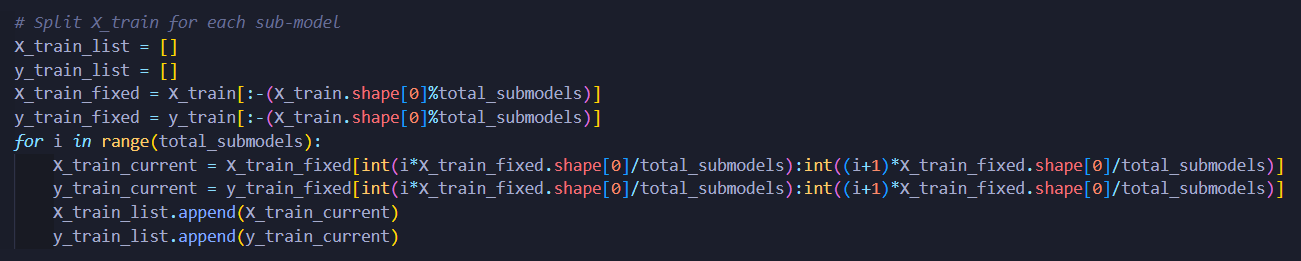


المن با 5 زیرمدل

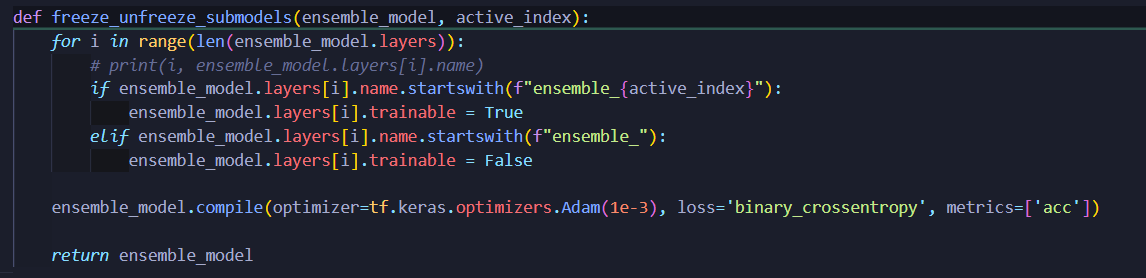


در نهایت به بخش آخر تمرین می­رسیم. در این بخش می­خواهیم با ترکیب زیر­مدل های **جردن** و **المن**، بهترین **مدل** **گروهی** ممکن را بسازیم. همچنین در این بخش با تقسیم بندی مجموعه داده آموزش به تعدادی مجموعه داده کوچکتر هر مجموعه داده کوچک را به یکی از زیرمدل هامی­دهیم. با این کار زیرمدل ها به صورت مستقل روی دیتاست های مجزا آموزش میبینند. البته این رویکرد ممکن است بهترین رویکرد نباشد و شاید بهتر بود مانند روشهای Boosting اشتباهات مدل اول را بوست کنیم و به مدل دوم بدهیم و الی آخر. با این کار زیر مدل ها به نوعی مکمل هم می­شدند.

با این حال به پیاده سازی این بخش از سوال میپردازیم. ابتدا مجموعه داده را به تعداد زیرمدل های جردن و المن تقسیم میکنیم.



سپس هر کدام از این زیرمجموعه ها را به عنوان ورودی یکی از زیرمدلها در نظر میگیریم. باید توجه کنیم که هنگامی که یک زیرمدل در حال آموزش دیدن است باید وزن سایر زیرمدل ها را غیرقابل آموزش(فریز) کنیم. این کار را با توجه به اسمی که برای هر لایه در نظر گرفتیم انجام می­دهیم.



سپس به آموزش دادن مدل نهایی میپردازیم و با آزمایش مقادیر مختلف سعی میکنیم بهترین مدل گروهی ممکن را پیدا کنیم. در جدول زیر تعدادی از آزمایش­های مربوط به این مدل گروهی را مشاهده میکنید. باید در نظر داشته باشیم که در این حالت افزودن زیرمدل های بیشتر به معنی کاهش داده آموزش هر کدام از زیرمدل هاست که این ممکن است باعث شود حتی از حالت بدون تقسیم داده ها نیز نتیجه بدتری بدست بیاوریم. اگر داده آموزش بیشتری داشتیم ممکن بود بتوانیم با این روش به نتایج بهتری برسیم اما در این مورد این شبکه گروهی نتیجه **بدتری** نسبت به شبکه های گروهی **بدون تقسیم داده­ها** داد.

|  |  |
| --- | --- |
| نتایج آزمایش­های شبکه گروهی **ترکیبی** با داده های مجزا | |
| صحت Test | توضیحات مدل |
| 0.6501 | [10, 75] Elman + [10, 75] Jordan |
| **0.71** | **[50, 75] Elman + [50] Jordan** |
| 0.688 | [75] Elman + [75] Jordan |

منابع:

<https://machinelearningmastery.com/stacking-ensemble-for-deep-learning-neural-networks/>

<https://www.tensorflow.org/guide/keras/custom_layers_and_models>

<https://www.tensorflow.org/guide/keras/custom_layers_and_models#the_layer_class_the_combination_of_state_weights_and_some_computation>

<https://datascience.stackexchange.com/questions/82416/difference-between-jordan-elman-and-normal-rnn>

<https://www.youtube.com/watch?v=e2sGq_vI41s&list=PLC112AD1C69432FDB&index=4>

<https://www.data-blogger.com/elman-rnn-implementation-in-tensorflow/>

**پایان**