

به نام خدا

گزارش کار پروژه شبکه عصبی درس یادگیری ماشین

حسین محمدی

۴۰۰۲۶۸۲۴

زمستان ۱۴۰۰

در این پروژه مفهوم نورون به آن صورت وجود ندارد. در عوض این مفهوم در دل وزن های لایه ها نهفته شده است.

پروژه به نحوی پیاده سازی شده است که امکان هر تعداد لایه با هر تعداد نورون وجود دارد. مسئله چالش برانگیز در این پروژه محاسبه فرمول انتشار رو به عقب برای هر لایه است. که با توجه به فرمول بدست آمده برای شبکه ۳ لایه میتوان آن را برای هر تعداد لایه تعمیم داد.

$$\frac{\partial E}{\partial w^3} = (-e * f^{3'}) * (o^2)^T$$

$$\frac{\partial E}{\partial w^2} = - \left(\left[(w^3)^T * (e * f^{3'}) \right] * f^{2'} \right) * (o^1)^T$$

$$\frac{\partial E}{\partial w^1} = - \left(\left[(w^2)^T * \left(\left[(w^3)^T * (e * f^{3'}) \right] * f^{2'} \right) \right] * f^{1'} \right) * (x)^T$$

مبینیم که فرمول لایه دوم تعمیم یافته از فرمول لایه سوم و همچنین فرمول لایه اول تعمیم یافته از فرمول لایه دوم میباشد.

در واقع مهم ترین موضوع سازگاری اندازه ماتریس ها با هم برای محاسبه بوده است.

هر لایه مقدار وزن خود، وزن جدید خود، مقدار net و مقدار خروجی خود را نگهداری میکند. در هربار استنتاج با شبکه عصبی، همه مقادیر بدست آمده برای محاسبه الگوریتم پس انتشار خطا ذخیره میشوند. پس از اتمام محاسبه وزن های جدید و ذخیره سازی، آنها را جایگزین وزن های فعلی کرده و استنتاج بعدی را شروع میکنیم.

در محاسبه وزن های جدید، نرخ یادگیری متناسب با اینکه در کدام epoch هستیم کم میشود تا شبکه همگرا شود. برای کم کردن از فرمول $new_lr = lr - 1 / epoch$ استفاده میکنیم.

در واقع شرط خاتمه الگوریتم تعداد epoch است که از قبل تعیین میکنیم.

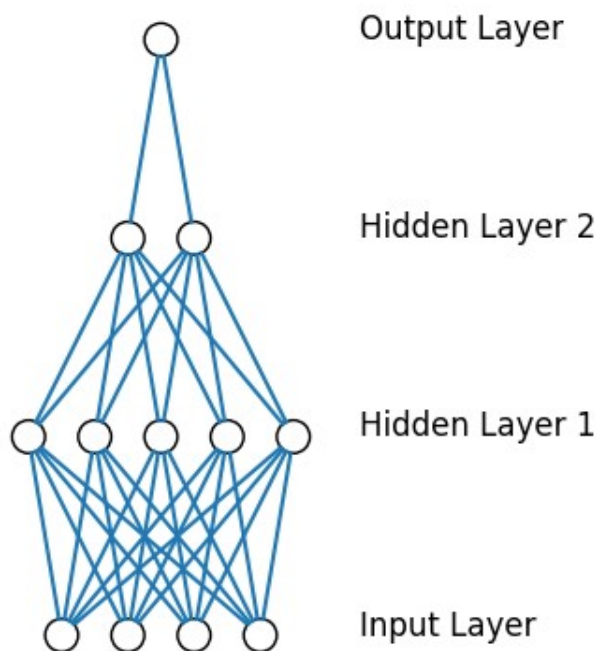
برای آموزش شبکه نیز هربار داده ورودی را به صورت تصادفی به دو بخش با ۷۰ درصد و ۳۰ درصد داده تقسیم میکنیم. قسمت بزرگتر برای آموزش و قسمت کوچکتر برای تست.

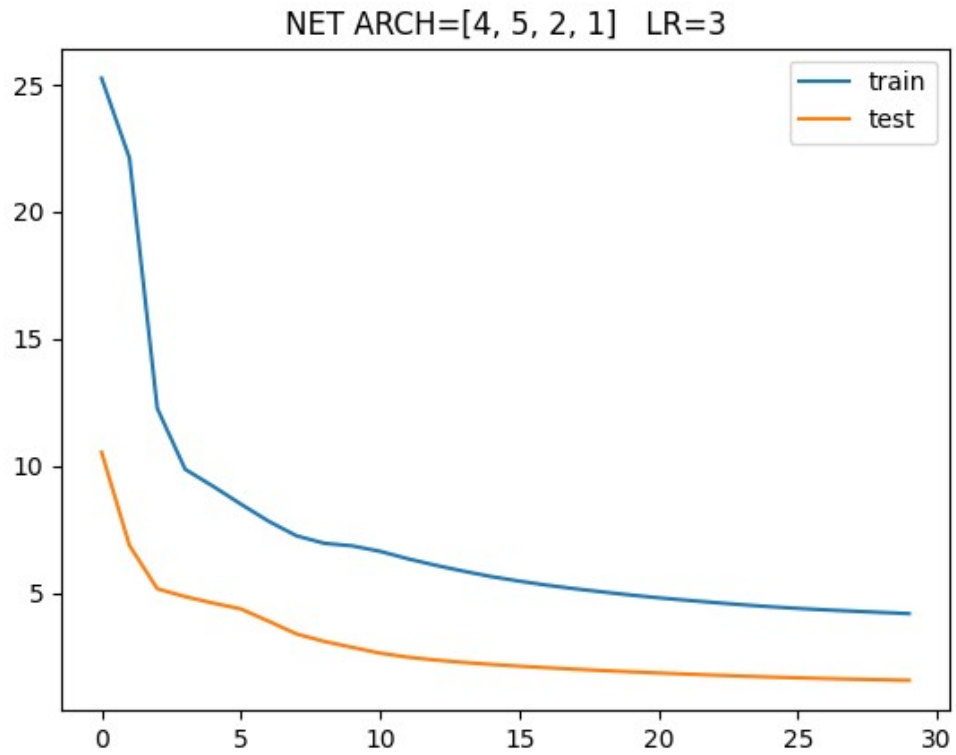
در هر دور قدرمطلق فاصله هر خروجی از خروجی مطلوب حساب شده و به عنوان خطا در نظر میگیریم. این خطا هم برای داده تست و هم برای داده آموزش محاسبه میشود.

شایان ذکر است در اینجا کلاس هر خروجی بر اساس نزدیک ترین مرکز تعیین میگردد نه بر اساس اینکه در کدام بازه قرار میگیرد که در واقعیت فرقی با حالت گفته شده در تعریف پروژه ندارد.

در نهایت خروجی های برنامه را مشاهده میکنیم.

Neural Network architecture





Last epoch confusion matrix

	Iris-setosa	Iris-versicolor	Iris-virginica
Iris-setosa	12	0	0
Iris-versicolor	0	13	1
Iris-virginica	0	0	19