

به نام خدا



پروژه ی نهایی درس هوش مصنوعی و محاسبات زیستی

تهیه کننده : حامد آجورلو

۹۷۱۰۱۱۶۷

استاد : دکتر حاجی پور

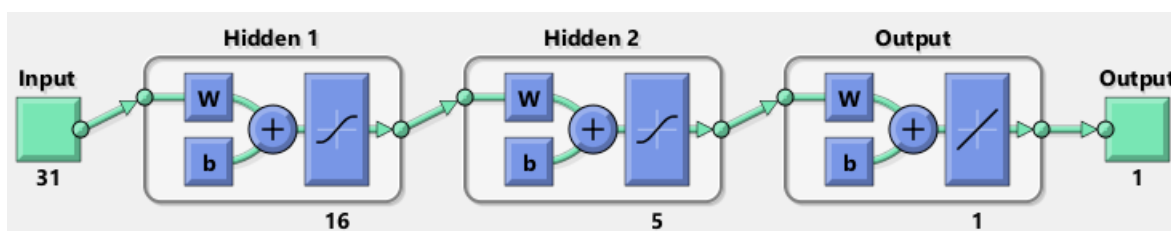
زمستان ۱۴۰۰

فاز ۱

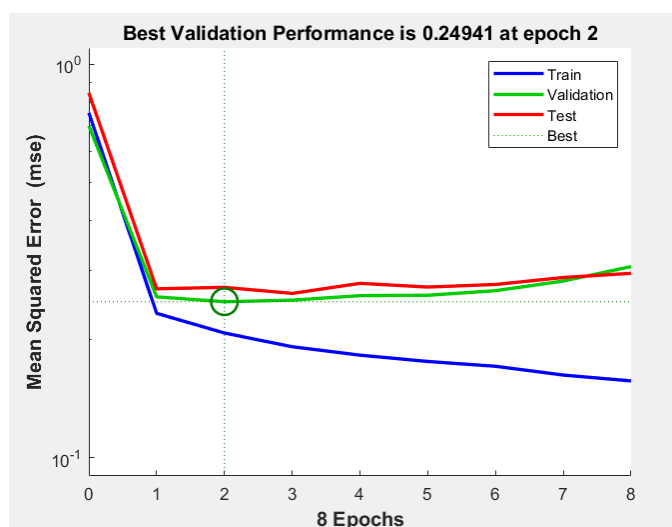
ابتدا داده ها را ارزیابی می نماییم و تعدادی از ویژگی های ذکر شده را انتخاب می نماییم .

بنده واریانس هر کانال ، کورلیشن بین هر دو کانال ، فرم فاکتور مختص هر کانال ، فرکانس بیشینه و فرکانس میانگین هر کانال را به عنوان ویژگی انتخاب کردم و بعد از محاسبه ی هر یک از ویژگی ها ، آن را در ماتریس feature ذخیره نمودم .

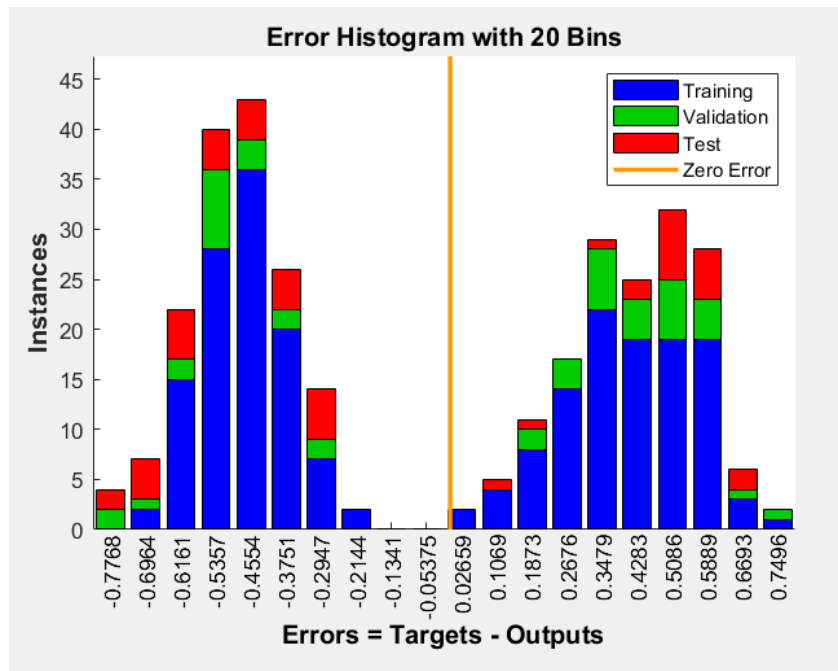
در مجموع ۸۹۶ ویژگی استخراج شد که در مرحله ی بعد با استفاده از معیار فیشر ، نمره ی هر یک از فیچر ها را محاسبه کردم و با آستانه گذاری 31 تا از موثر ترین ویژگی ها را انتخاب نمودم. با در نظر گرفتن این ویژگی ها شبکه ای از نوع MLP آموزش داده شد و حالات مختلف برای تعداد نوروں های لایه پنهان و تعداد لایه ها آزمایش شد و با روش 5fold cross validation دقت هر یک از شبکه ها سنجیده شد و در نهایت شبکه ای با دو لایه ی پنهان که به ترتیب ۱۶ و ۵ نوروں دارند طراحی شد ، تصویر شماتیک شبکه MLP آموزش داده شده را در شکل زیر مشاهده می فرمایید :



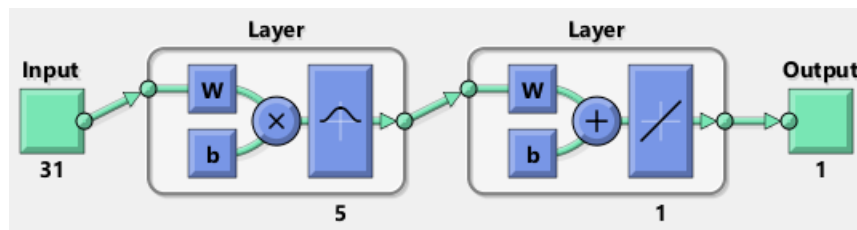
پرفورمنس شبکه ی آموزش داده شده :



هیستوگرام خطا :



با استفاده از همین ویژگی ها یک شبکه ی RBF طراحی می نماییم و از حالت های مختلف تست شده ، شبکه ای که بهترین دقت را بر روی داده ی اعتبارسنجی بدست می آورد را انتخاب می نماییم. شبکه ی بدست آمده به قرار زیر است :



خطای شبکه با معیار MSE :

$$MSE = 0.249997$$

ملاحظه می فرمایید که تقریباً هر دو شبکه پرفورمنس یکسانی دارند و خطای MSE تقریباً برابری دارند ولی شبکه MLP مقدار خیلی کمی بهتر است.

ویژگی های انتخاب شده عبارتند از واریانس کانال ۲۷ ام و کورلیشن تعدادی از کانال ها که اندیس ویژگی مربوطه در بردار eff_i ذخیره شده است.

در این بخش ، از ۸۹۶ ویژگی ابتدایی ، برای کاهش زمان اجرای الگوریتم تکاملی ۵۰ ویژگی برتر را انتخاب می نماییم و روش رمز گذاری را به اینگونه تعریف می کنیم :

یک کروموزم با ۵۰ ژن به صورت یک آرایه با ۵۰ درایه که از صفر و یک به وجود آمده است.

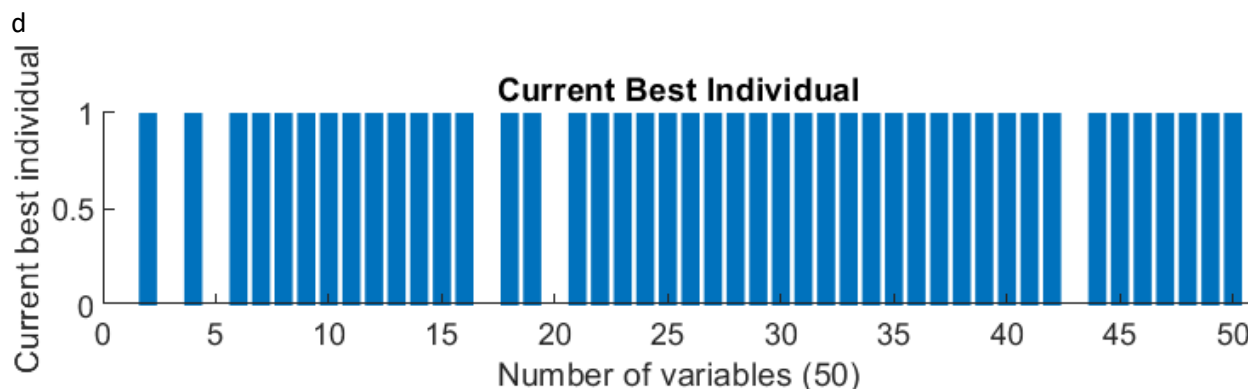
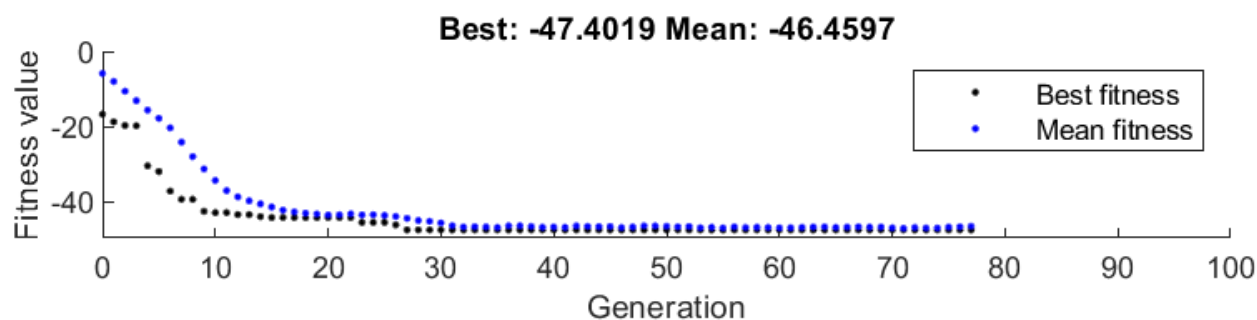
0	1	0	1	1	0	0	1	1
---	---	---	---	---	---	---	---	---

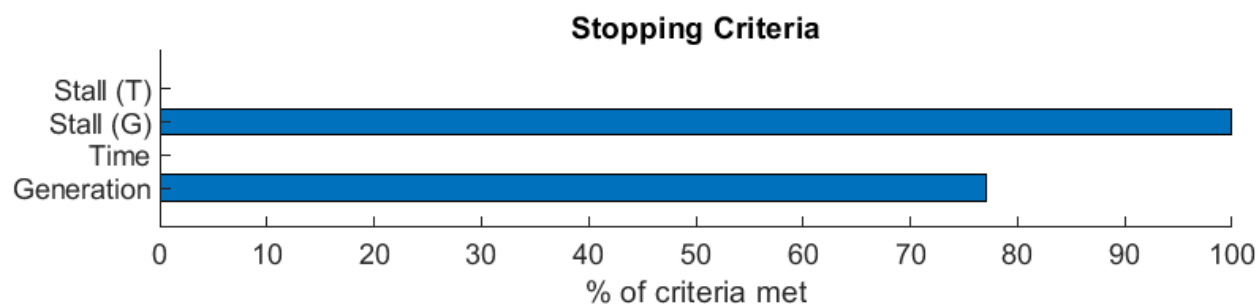
$1 \times N$

یک بودن هر یک از ژن ها نشانگر حضور فیچر نام در جمعیت مربوطه می باشد .

تابع سازگاری را مبتنی بر فرمول فیشر گروهی تعریف می کنیم که مناسب بودن هر دسته از ویژگی ها را نشان دهد. جمعیت اولیه را به صورت رشته ی بیتی ، کراس اور یک نقطه ای و جهش یونیفرم را انتخاب می نماییم و شرط خاتمه ی شبیه سازی را رسیدن به ۱۰۰ نسل قرار می دهیم.

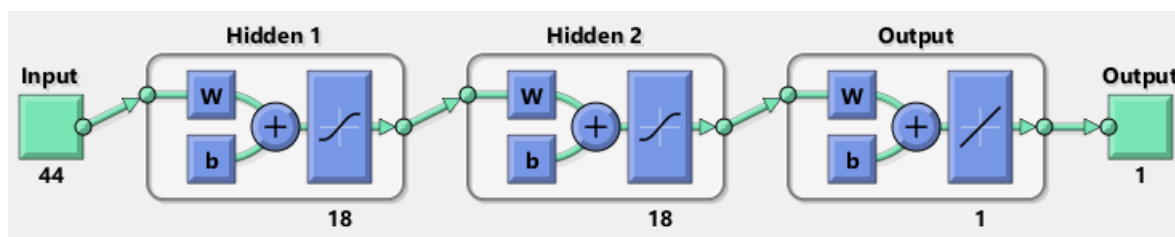
با کمک optimtool شبیه سازی را انجام داده و نتایج زیر را بدست می آوریم :



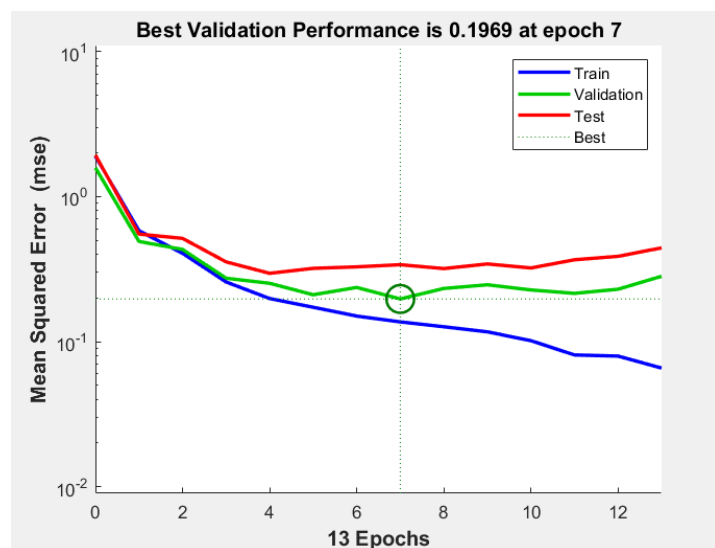


حال به کمک ویژگی های بدست آمده شبکه ی MLP را مشابه آنچه در فاز ۱ انجام دادیم بدست می آوریم :

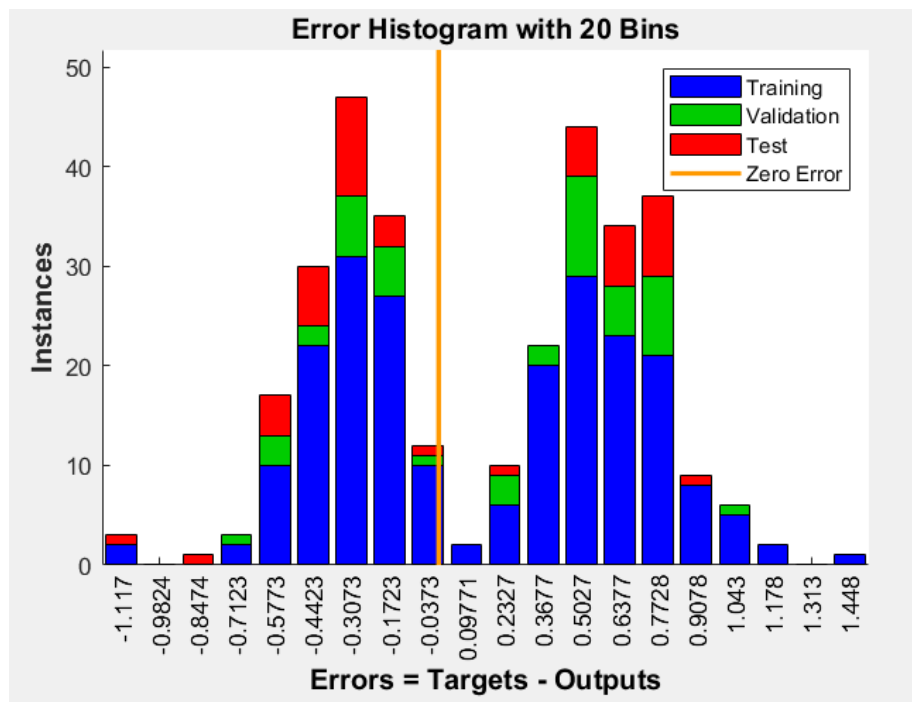
این شبکه بهترین پرفورمنس را بر روی داده ی اعتبارسنجی در الگوریتم 5fold cross validation داشته است :



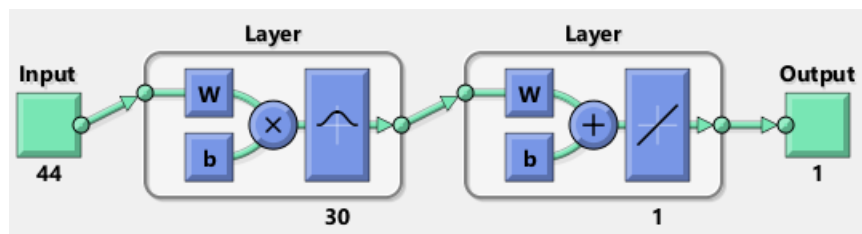
پرفورمنس شبکه MLP آموزش داده شده :



هیستوگرام خطا :



با استفاده از همین ویژگی ها یک شبکه ی RBF طراحی می نماییم و از حالت های مختلف تست شده ، شبکه ای که بهترین دقت را بر روی داده ی اعتبارسنجی بدست می آورد را انتخاب می نماییم. شبکه ی بدست آمده به قرار زیر است :



خطای شبکه با معیار MSE :

$$MSE = 0.249997$$

بین دو شبکه ی طراحی شده ، شبکه MLP پرفورمنس بهتری دارد. بنابر معیار MSE لیبل های بدست آمده از فاز های ۱ و ۲ در بردارهای Predicted در هر فاز ذخیره شده اند.