به نام خدا



دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

درس شبکه‌های عصبی

استاد صفابخش

تمرین اول

علیرضا مازوچی

۴۰۰۱۳۱۰۷۵

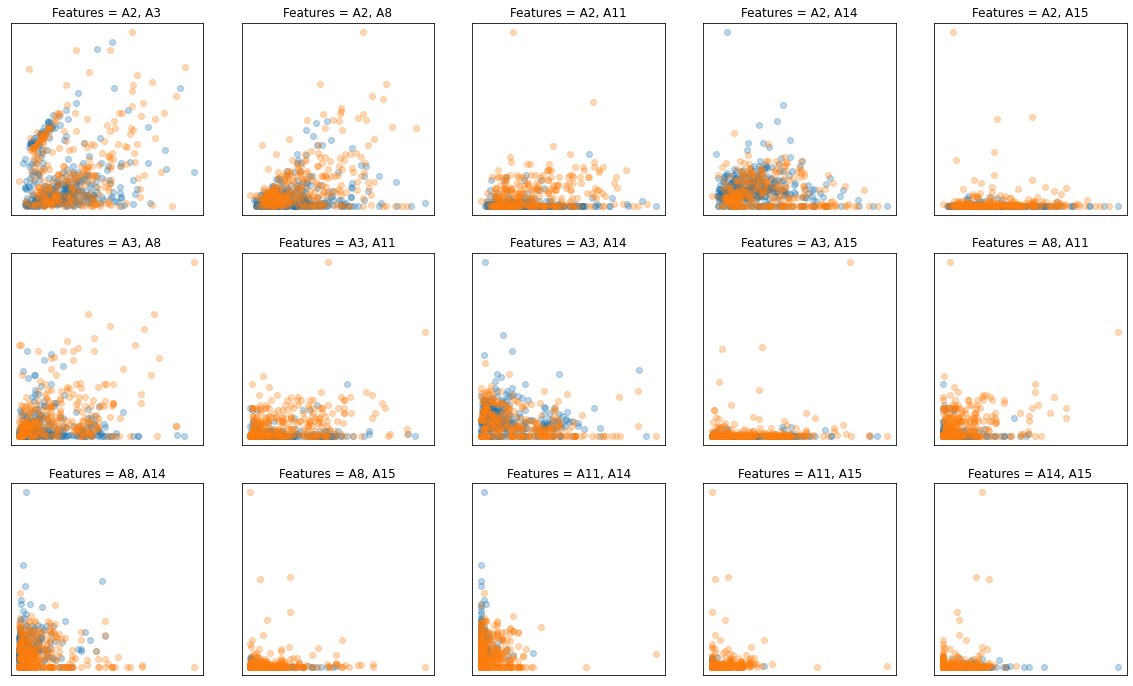
سوال ۱

مطابق با توضیحات موجود در مجموعه‌داده در حدود ۵ درصد داده‌ها یک یا چند ویژگی دارای مقدار تهی است. برای ستون‌هایی که دارای داده‌های پیوسته هستند از میانگین سایر داده‌ها و برای ستون‌هایی که دارای مقادیر گسسته هستند از ماکسیمم مقدار برای پر کردن مقدار از دست رفته بهره می‌گیریم. برای تمام ستون‌های ویژگی یک نرمال‌سازی مطابق با فرمول زیر هم به کار می‌گیریم تا رنج مقادیر برای تمام این ستون‌های مشابه با یکدیگر باشد:

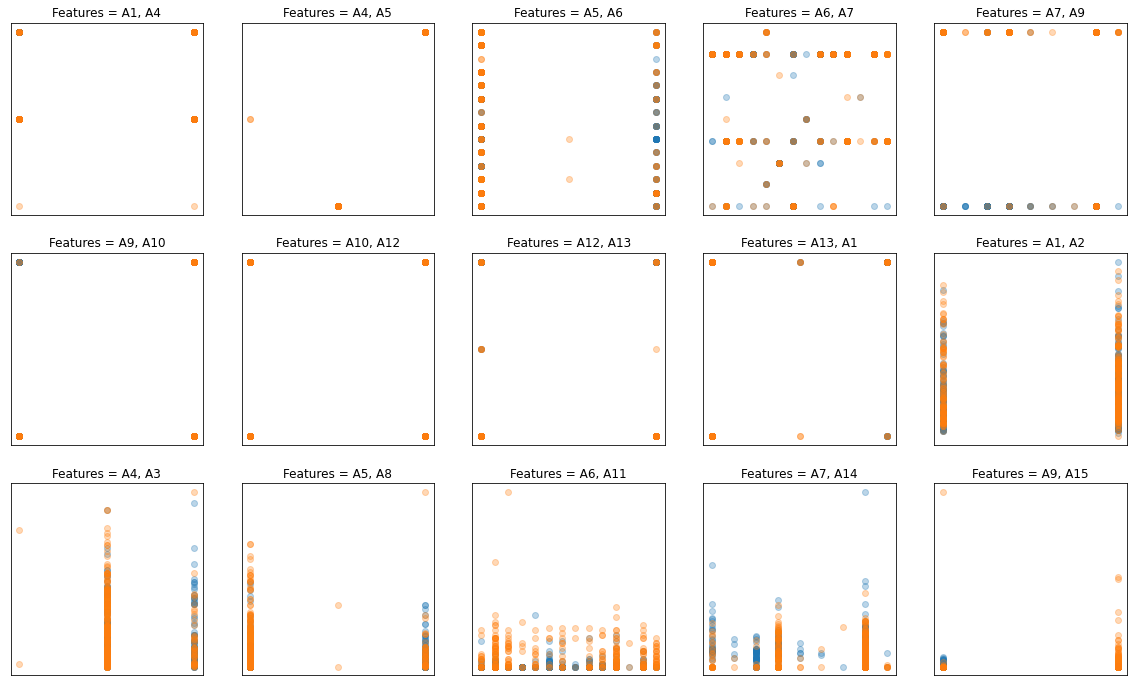
نهایتا مقادیر + و - ستون کلاس را به مقدار ۰ و ۱ تبدیل کردیم.

سوال ۲

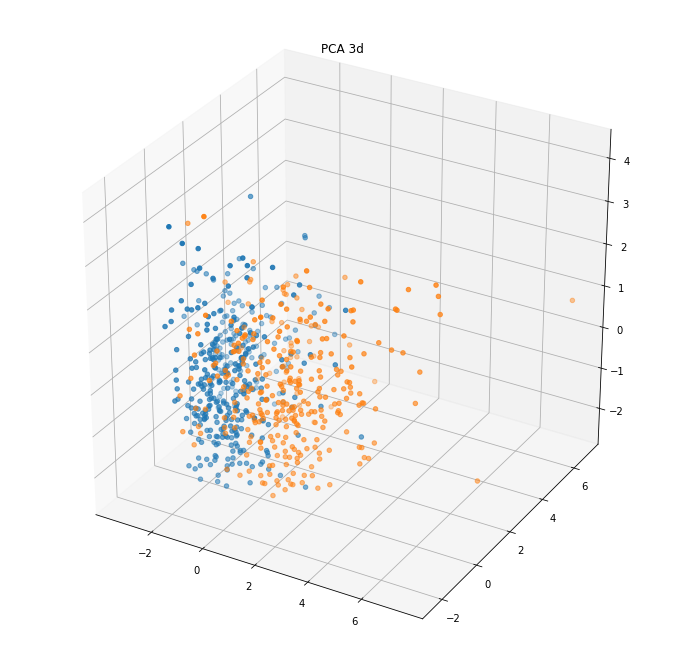
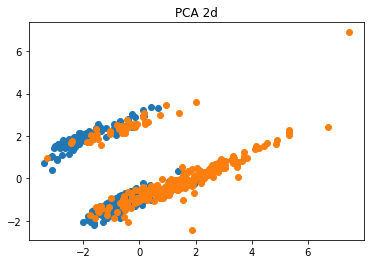
برای بررسی این مورد، چندین بررسی را انجام دادیم. اول آنکه ویژگی‌ها پیوسته را به صورت دو به دو در یک نمودار ترسیم کردیم ولی در هیچ کدام از نمودار‌ها داده‌های دو کلاس از یکدیگر به صورت خطی جدا نشد:



ویژگی‌های پیوسته با توجه به رنج وسیع‌تری که نسبت به ویژگی‌های گسسته دارند احتمال بیشتری برای جداکردن کلاس‌ها دارند ولی با این حال ممکن است ویژگی‌های گسسته داده‌‌ها را جداکنند. لذا تعدادی از ترکیب‌های ویژگی‌ها گسسته با هم و با ویژگی‌های پیوسته را بررسی کردیم ولی در این حالت هم موفقیتی حاصل نشد:



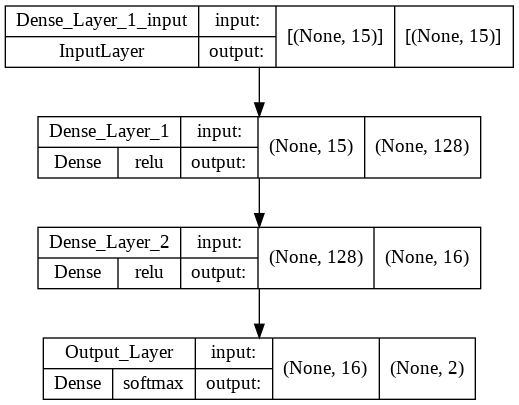
بار دیگر از روش PCA استفاده کردیم ولی این بار هم جداسازی داده‌ها انجام نشد:



حال که این روش‌های اولیه جواب ندادند،‌ نوبت به بررسی یک روش پیچیده ولی با جواب قطعی می‌رسد. این بار از روش SVM با حاشیه سخت استفاده می‌کنم. اگر بتوان یک مدل SVM‌ با حاشیه سخت پیدا کرد که به دقت ۱۰۰٪ برسد یعنی داده‌ها خطی جداپذیرند و اگر چنین مدلی وجود نداشته باشد یعنی خطی جداپذیر نیستند. در بررسی‌ای که من انجام دادم چنین مدلی پیدا نشد و لذا داده‌ها خطی جدا ناپذیرند.

سوال ۳

به عنوان اولین تلاش شبکه عصبی زیر را ایجاد کردم. در این شبکه لایه ورودی شامل ۱۵ ویژگی است. سپس از این ۱۵ ویژگی توسط یک لایه Dense ۱۲۸ ویژگی استخراج می‌شود. در لایه Dense بعد ۱۶ ویژگی ترکیبی و پیچیده حاصل می‌شود و نهایتا در لایه خروجی دو ویژگی ایجاد می‌شود که هر کدام متناسب با احتمال تعلق داده به یکی از دو کلاس موجود است. در لایه‌های Dense میانی از ReLU به عنوان تابع فعال‌سازی استفاده شده است تا عملکرد غیرخطی به مدل داده شود و در آخرین لایه که لایه خروجی باشد از یک لایه Softmax استفاده شده است تا خروجی از جنس احتمال باشد و با Sparse Categorical Cross Entropy که به عنوان تابع خطا استفاده می‌شود سازگار باشد.



بعد از ۲۰ گام آموزش به صحت ۹۲٪ روی داده‌های آموزش و ۸۷٪ روی داده‌های اعتبارسنجی رسیدم.