به نام خدا



دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

درس شناسایی آماری الگو

استاد رحمتی

تمرین اول

علیرضا مازوچی

۴۰۰۱۳۱۰۷۵

سوال ۱

پیش از جواب‌دادن به سوالات لازم است تا روش طراحی سامانه را تبیین کنم. من قصد دارم تا سه دسته‌بند طراحی کنم و سپس این سه دسته‌بند در یک چهارچوب گروهی (Ensemble) جواب نهایی را برگردانند. دسته‌بند اول با چهره شخص و یک شبکه عصبی تشخیص می‌دهد که فرد با چه احتمالی در بازه مناسب قرار دارد. دسته‌بند دوم با پردازش صوت کودک با یک شبکه عصبی دیگر تشخیص می‌دهد که کودک در بازه سنی مناسب قرار دارد یا نه؟ دسته‌بند سوم با ویژگی قد و وزن کودک احتمال تعلق داشتن کودک به بازه سنی مناسب را بیان می‌کند. نهایتا خروجی هر سه دسته‌بند با به یک MLP‌ ساده می‌دهیم تا نتیجه نهایی را اعلام کند.

الف) مسئله از نوع دسته‌بندی یا Classification است؛ چراکه قرار است کل بازدیدکنندگان را به دو کلاس گروه سنی مناسب و گروه سنی نامناسب تقسیم کنیم.

ب) به سنسور تشخیص چهره،‌ صوت، قد و وزن نیاز داریم.

ج) ما یک مجموعه‌داده نیاز داریم که برای هر داده یک تصویر از تمام‌رخ چهره، یک صوت بعد از خواندن یک متن، قد و وزن را در خود داشته باشد.

د) می‌توان داده‌ها را از بازدیدکنندگان فعلی تا زمانی استقرار سامانه تهیه کرد. یعنی هر تصویر و سایرمشخصات هر کودک بازدیدکننده را می‌توان برداشت و یک ناظر انسانی مسئول تعیین کلاس کودک می‌شود. باتوجه به آمار بالای بازدید پارک دیزنی، با این شیوه می‌توان در مدت زمان قابل قبولی حجم مناسب داده جمع‌آوری شود. البته باید توجه داشت ممکن است قوانین حفظ حریم خصوصی مانع از جمع‌آوری داده تمام بازدیدکننده‌ها بدون اجازه آن‌ها باشد که در این صورت لازم است این جمع‌آوری داوطلبانه و با آگاهی بازدیدکننده‌ها صورت بگیرد.

علاوه بر شیوه فوق می‌توان از مجموعه‌داده‌های آماده هم کمک گرفت که البته هر کدام بخشی از ویژگی‌ها مدنظر ما را دارند و می‌توانند در توسعه زیرمدل‌های ما مورد استفاده قرار بگیرند.

ه) ویژگی قد، وزن،‌ یک تصویر که یک آرایه دوبعدی از ویژگی‌هاست و صوت که یک آرایه یک‌بعدی از ویژگی‌هاست.

و) برای ویژگی قد و وزن پیش‌پردازش خاصی نیاز نیست. در مورد صوت، لازم است تا پیش‌پردازشی برای حذف نویز محیط و سایر نویز‌های این حوزه انجام گیرد. در مورد تصویر هم باید پیش‌پردازش‌هایی برای شفاف کردن عکس، حذف لرزش موقع تصویر و تاری عکس و مواردی از این دست صورت بگیرد.

ز) مهم‌ترین چالش جمع‌آوری داده است. باتوجه به پیچیدگی مدل و دارابودن انواع ویژگی‌ها نیاز به تعداد داده بالاست. قوانین حفظ حریم خصوصی هم ممکن است کار را سخت‌تر بکنند. در این حالت ممکن است داده به اندازه جمع‌آوری نشود که مجبور خواهیم بود از مدل‌های ساده برای بخش تصویر و صوت استفاده کنیم که بی‌شک در دقت مدل موثر است. حتی این احتمال وجود دارد که مجبور باشیم بخشی از مدل را مثلا بخش تصویر را حذف کنیم.

یکی دیگر از چالش‌های سیستم جدید،‌ امکان تغلب در آن است. درحالی که ویژگی‌هایی مانند قد به سختی مورد تغلب واقع می‌شود ولی در مورد صوت و تصویر ممکن است کاربر تغییراتی در صدا و چهره خود بدهد که سیستم را به شک بیاندازد. در مورد وزن هم امکان افزایش آن با لباس‌های سنگین وجود دارد. همچنین به دلیل کرونا ( و بیماری‌های آتی!) ممکن است امکان دریافت چهره کامل به سختی صورت بگیرد.

ح) نقطه قوت اصلی سیستم، دقت بیشتر و توجه به جوانب مختلف نسبت به سیستم قبلی است. در سیستم قبلی تنها به قد توجه می‌شد که ممکن بود یک کودک که قد آن نسبت سن واقعی‌اش متداول نباشد را به خطا بیاندازد درحالی که در این سیستم مبتنی بر چندین زیرمدل است. نقطه قوت دیگری که شاید بتوان برای آن درنظر گرفت آن است که خروجی سیستم فعلی برای کودکان و والدینشان پذیرفتنی‌تر است. در سیستم قبلی محتمل بود تا برخی از کودکان و والدین توجه تنها به یک خط‌کش را عادلانه ندانند درحالی که در سیستم فعلی و پس از جمع‌آوری اطلاعات مختلف و نهایتا تصمیم‌گیری توسط کامپیوتر امکان متقاعدکردن بهتری وجود خواهد داشت.

درکنار نقاط قوت، نقاط ضعفی وجود دارد. سیستم فعلی نسبت به سیستم قبل، پیچده‌تر است و برای استفاده از آن هم نیاز به تجهیزات مختلف سخت‌افزاری و نرم‌افزاری است. استفاده از سیستم پیشنهادی وقت کاربران را می‌گیرد چراکه آن‌ها باید چهار تست را به جای یک تست بدهند. اگرچه می‌توان این‌ها را باهم گرفت. (مثلا در حین ایستادن بر روی وزنه، قد و تصویر هم برداشته شود و در همین حین از کودک خواسته شود یک متن یا شعر را بخواند.) همچنین جمع‌آوری داده‌ها همانطور که گفته شد سخت، زمان‌بر و پرهزینه است.

سوال ۲

الف)

ب)

ج)

د)

ه) تعداد گوشه،

و) تعداد گوشه،

ز) تعداد گوشه،

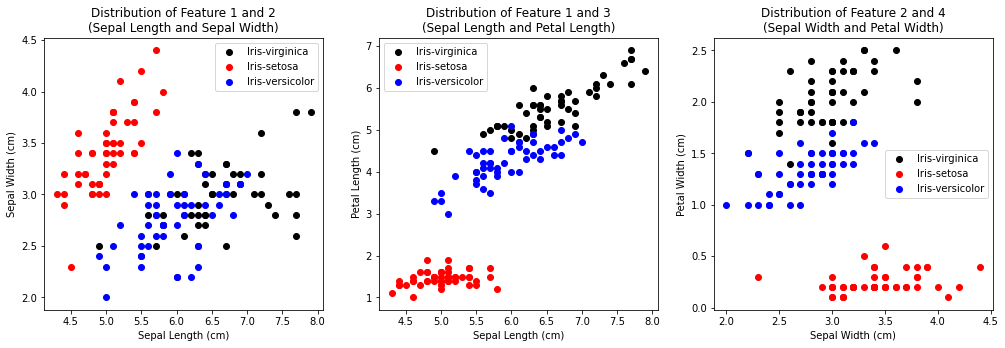
ح) تعداد گوشه،

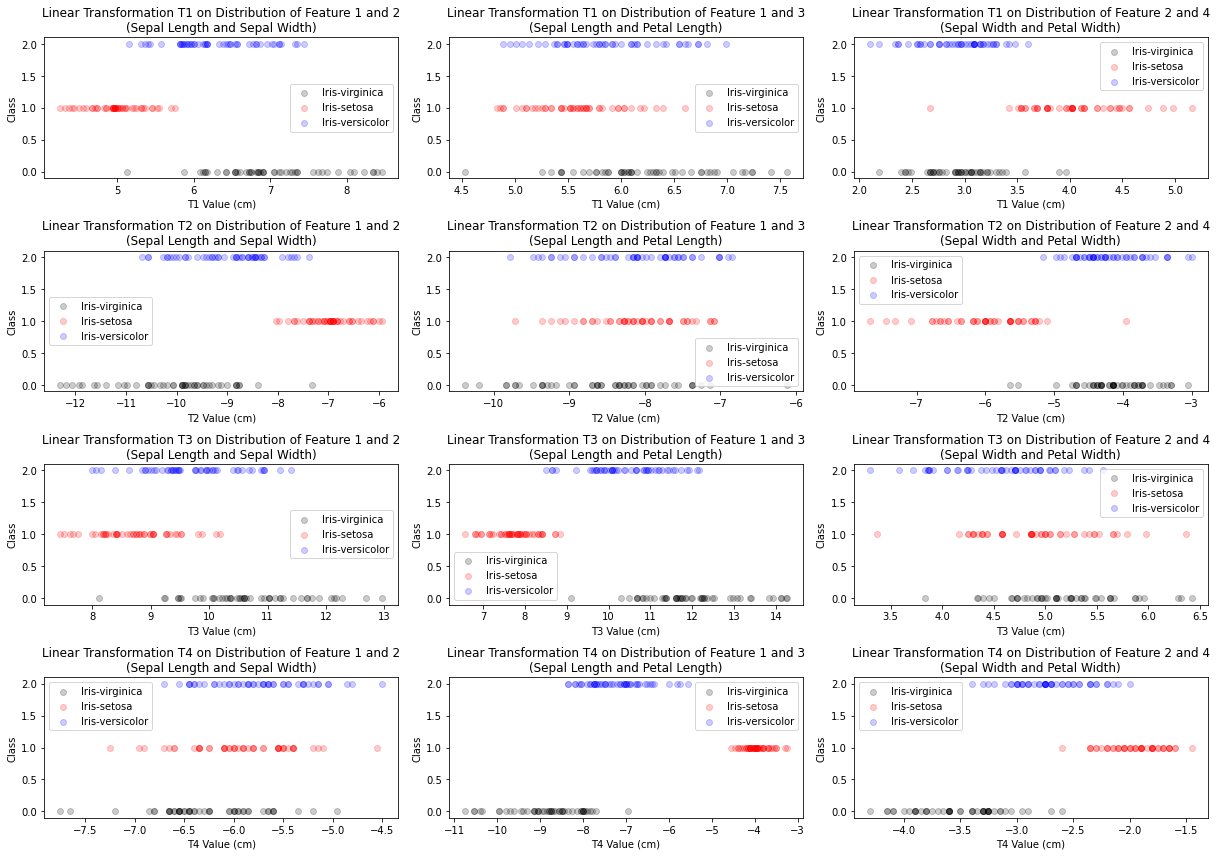
ط) تعداد گوشه،

ی) تعداد گوشه،

سوال ۳

الف)



ب)

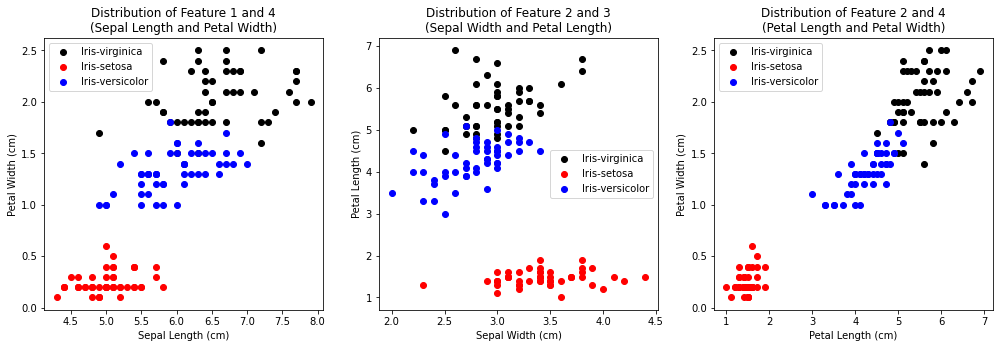
ج)

برای ویژگی ۱ و ۲، از همه مناسب‌تر به نظر می‌رسد اگرچه هم قابل قبول است. دو تبدیل و نتوانسته‌اند برخلاف دو تبدیل دیگر، کلاس نسبتا مجزای قرمز را به خوبی جداکنند. در مورد و هر دو تقریبا مشابه هستند اما تبدیل کمی کلاس قرمز را بهتر جدا کرده است.

برای ویژگی ۱ و ۳، مناسب‌ترین است. در تبدیل و داده‌های هر سه کلاس روی هم افتاده است. در تبدیل هم دو کلاس آبی و سیاه همپوشانی زیادی دارند و داده‌های کلاس قرمز نسبت به پراکنده‌تر است.

برای ویژگی ۲ و ۴، مناسب‌ترین است.‌ تبدیل داده‌های هر سه کلاس را در یک نقطه نگاشت می‌کند. و دو کلاس آبی و سیاه را روی هم می‌اندازد درحالی که در سه کلاس تاحد قابل قبولی از هم جدا شده‌اند.

د)



ه) با توجه به نمودار‌های قسمت د می‌توان حدس زد که برای جفت ویژگی (۱ و ۴) و (۲ و ۳) خطوط موازی با محور x‌ها با دقت بسیار خوبی می‌توانند داده‌های کلاس‌های مختلف را از هم جدا کنند. لذا برای این دو قسمت از تبدیل خطی ]۱ و ۰ [ استفاده کردم. برای جفت ویژگی ( ۲ و ۴ ) شاید خطوط مورب 45 درجه ( بدون توجه به مقیاس) به یکی از بهترین نتایج ممکن ختم شود. نهایتا با اختلاف مقیاس دوبرابر ویژگی اول و دوم، از تبدیل خطی ]۲ و ۱ [ استفاده کردم:

