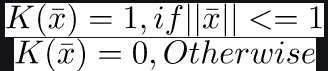
**تمرین 3**

**سوال1**

کرنل روشی است که برای گرفتن داده به عنوان ورودی و تبدیل به فرم مورد نیاز پردازش داده استفاده می شود. "کرنل " به دلیل مجموعه ای از توابع ریاضی مورد استفاده در SVM پنجره ای را برای دستکاری داده ها به کار می برد. بنابراین ، عملکرد کرنل به طور کلی مجموعه آموزش داده ها را به گونه ای تغییر می دهد که سطح تصمیم گیری غیر خطی قادر به تبدیل شدن به یک معادله خطی در تعداد بیشتری از فضاهای بعد باشد.

معادله عملکرد استاندارد کرنل:



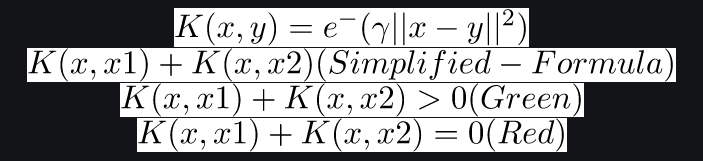
توابع اصلی کرنل:

کرنل گوسی: برای ترنسفورم کردن داده استفاده می شود ، درصورتی که اطلاعات قبلی در مورد داده وجود نداشته باشد.



RBF:

همان عملکرد کرنل گوسی است ، به همراه افزودن روش radial basis برای بهبود ترنسفورمیشن.



کرنل سیگموئید:

این عملکرد معادل مدل دو لایه ای از شبکه عصبی پرسپترون است که به عنوان عملکرد فعال سازی برای نورونهای شبکه عصبی استفاده می شود.

**سوال7**

یکی از محبوب‌ترین و در عین حال ساده‌ترین الگوریتم‌های درخت‌های تصمیم، درخت تصمیمِ CART است که کاربردهای زیادی در طبقه بندی و رگرسیون دارد. CART که خود مخفف Classification and Regression Tree است بر اساس درخت های دودویی (باینری) بنا نهاده شده است. در این درس می‌خواهیم بیشتر با نحوه‌ی ساختِ درختِ CART آشنا شویم. این درخت (و البته درخت‌های دیگر) می‌تواند پایه‌ای برای الگوریتم‌های پیچیده‌تر مانند جنگلِ تصادفی (Random Forest) باشد.الگوریتمِ درختِ تصمیمِ CART برای ساختِ درختِ تصمیم، داده‌ها را به قسمت‌های دو‌تایی تقسیم کرده و بر اساس آن‌ها درخت دو‌دویی (باینری) را می‌سازد.

**الگوریتم ID3** یکی از الگوریتم‌های پایه برای ساختِ درخت‌های تصمیم است. لگوریتم ID3 وظیفه پیدا کردنِ ویژگی‌هایی دارای اطلاعات زیادتر (Gain بیشتر) را دارد و آن‌ها را در سطوحِ بالاتری از درخت قرار می‌دهد. هر بار که یک ویژگی در سطحی از درخت انتخاب شد، زیر درخت‌های آن نیز دقیقا به همان صورت (ویژگی‌هایی با اطلاعات بالا) انتخاب می‌شوند و در سطوح و گره‌های بعدی قرار می‌گیرند. البته وقتی یک گره از درخت انتخاب شد، برای ساختِ زیر درخت‌های دیگر، مجموعه داده‌ها بر اساس مقدارِ گره‌ی انتخاب شده در شاخه‌های بالاتر، کوچکتر می‌شوند و هر چه در درخت پایین‌تر می‌رویم (به برگ‌ها نزدیک‌تر می‌شویم)، مجموعه داده‌ها برای محاسبه‌ی مقدار اطلاعات کمتر می‌شوند.

**سوال 9**

عمق درخت که نشانگر تعداد گره‌ها از ریشه تا گره‌های پایانی است. درخت هرگاه به حداکثر عمق تعریف شده‌ای برسد و یا تعداد نمونه‌ها در گره جاری از مقدار تعریف شده‌ای کمتر باشد، متوقف می‌شود.

**سوال10**

هرس هست یک متراکم سازی داده ها تکنیک در فراگیری ماشین و الگوریتم های جستجو که اندازه را کاهش می دهد درختان تصمیم با برداشتن بخشهایی از درخت که غیرقابل انتقاد و اضافی هستند برای طبقه بندی موارد. هرس پیچیدگی فینال را کاهش می دهد طبقه بندی، و از این رو دقت پیش بینی را با کاهش از ، بهبود می بخشد نصب بیش از حد. درختی که بیش از حد بزرگ خطراتی را به همراه دارد نصب بیش از حد داده های آموزش و تعمیم ضعیف به نمونه های جدید. یک درخت کوچک ممکن است اطلاعات مهم ساختاری مربوط به فضای نمونه را به دست نیاورد. با این حال ، تشخیص زمان الگوریتم درخت دشوار است زیرا تشخیص اینکه آیا افزودن یک گره اضافی به طور چشمگیری خطا را کاهش می دهد ، غیرممکن است. این مشکل به عنوان شناخته می شود اثر افق. یک استراتژی معمول این است که درخت را رشد دهید تا زمانی که هر گره شامل تعداد کمی از موارد باشد و سپس از هرس برای حذف گره هایی استفاده کنید که اطلاعات بیشتری ارائه نمی دهند. هرس باید اندازه درخت یادگیرنده را کاهش دهد بدون اینکه دقت پیش بینی را که توسط a اندازه گیری می شود کاهش دهد اعتبار سنجی متقابل تنظیم. تکنیک های زیادی برای هرس درخت وجود دارد که در اندازه گیری که برای بهینه سازی عملکرد استفاده می شود متفاوت است.

**سوال12**

RF یا جنگل‌های تصمیم تصادفی (Random forest) یک روش یادگیری ترکیبی برای دسته‌بندی، رگرسیون می‌باشد، که بر اساس ساختاری متشکل از شمار بسیاری درخت تصمیم، بر روی زمان آموزش و خروجی کلاس‌ها (کلاس‌بندی) یا برای پیش‌بینی‌های هر درخت به شکل مجزا، کار می‌کنند. جنگل‌های تصادفی برای درختان تصمیم که در مجموعه آموزشی دچار بیش برازش می‌شوند، مناسب هستند. عملکرد جنگل تصادفی معمولا بهتر از درخت تصمیم است، اما این بهبود عملکرد تا حدی به نوع داده هم بستگی دارد.

**سوال 15**

بله امکان پذیر است، در واقع این روش ، یک روش کیفی برای تجزیه و تحلیل و پیش بینی در نظر گرفته می شود زیرا برای اجرای آن نیاز به مقدار کمی اطلاعات آماری است. در سالهای گذشته ، روش درختان تصمیم گیری نیز به طور مکرر در اثبات تصمیمات در زمینه اقتصادی اجتماعی آغاز شده است.