به نام خدا

عنوان:

تکلیف شماره هفتم درس یادگیری ماشین-SVM

استاد:

دکتر پدرام

دانشجو:

محمدعلی مجتهدسلیمانی

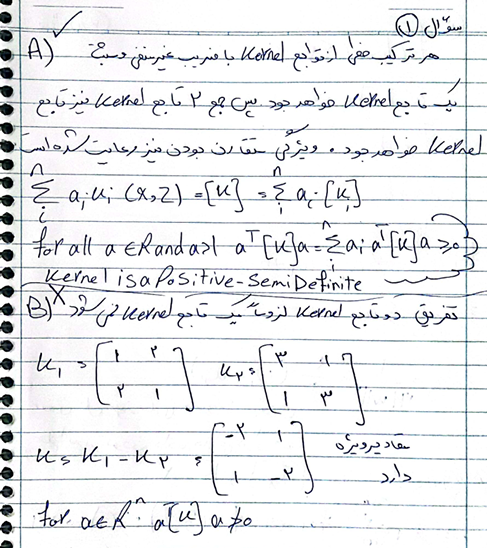
شماره دانشجویی:

4033904504

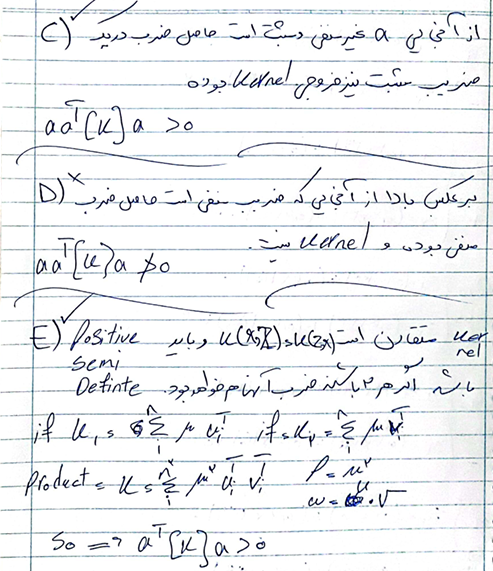
تاریخ:

20/8/1403

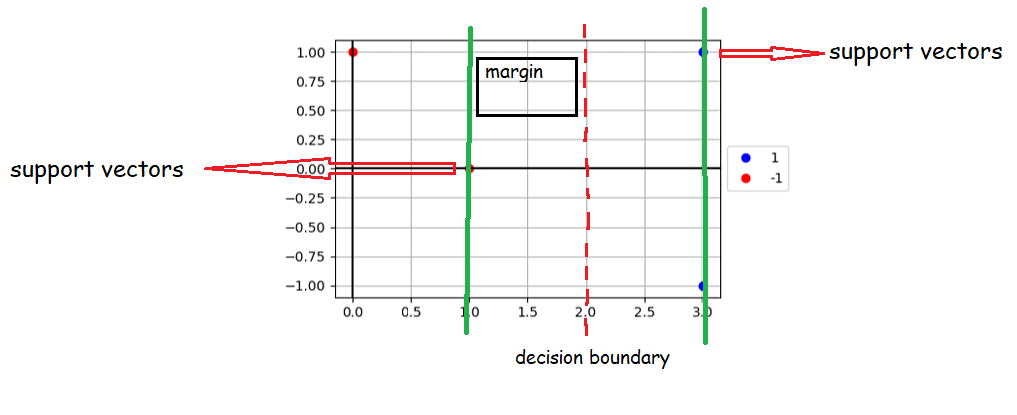
**سوال 1.**



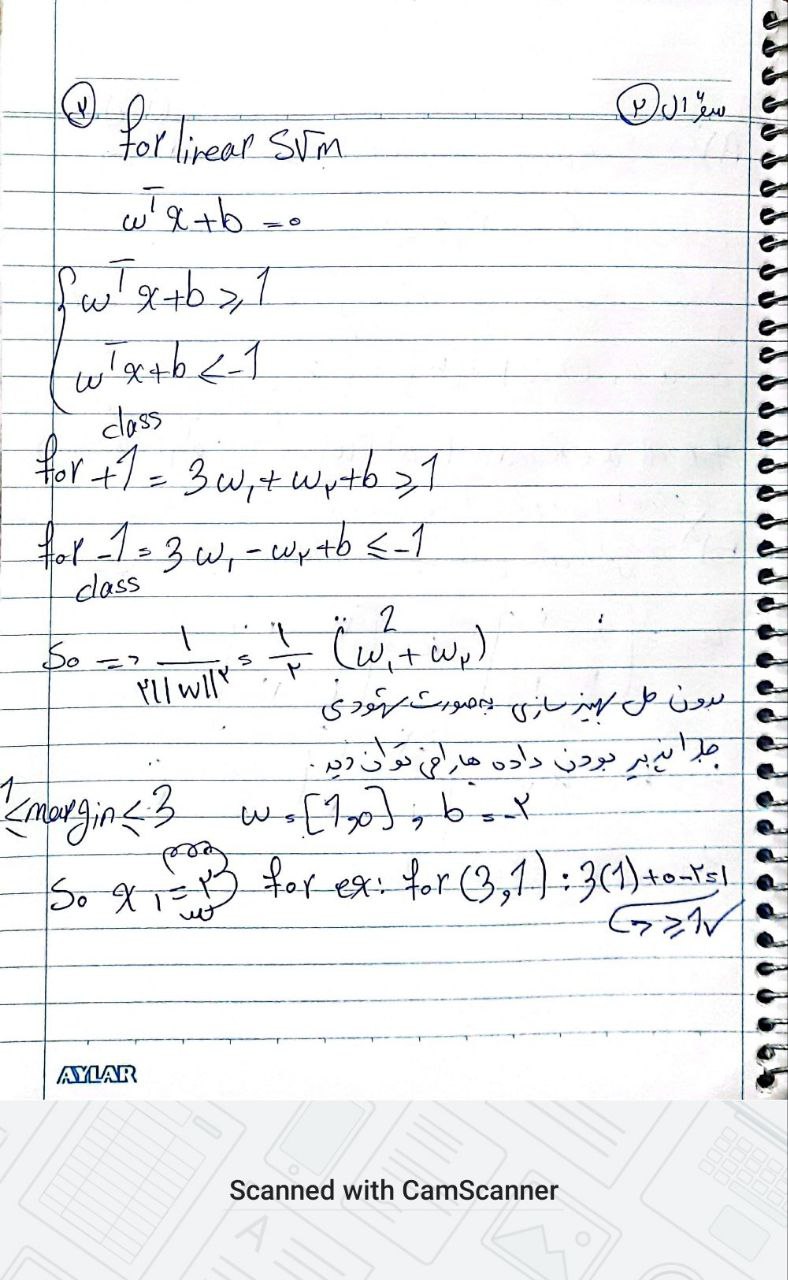
\*ادامه سوال در صفحه بعد قرار دارد.



**سوال 2.**



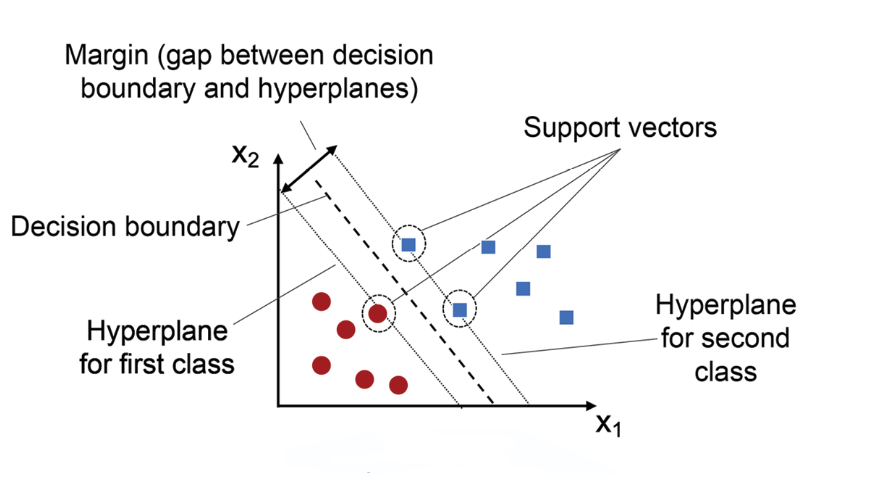
\*محاسبات در صفحه بعد قرار دارد.

**محاسبات:**  


**سوال 3.**

الف)

درست است. بعد از آموزش SVM تنها چیزی که برای خط تصمیم گیری (decision boundary) مهم هست بردار های پشتیبان (support vector) ها هستند و بقیه نقاط داده های آموزشی میتوانند بدون تاثیری در کارایی مدل حذف شوند. بخاطر اینکه فقط بردار های پشتیبان هستند که بیشترین margin را میتوانند مشخص کنند. margin به فاصله بین برادرهای پشتیبان و خط تصمیم گیری گفته میشود و بردار های پشتیبان به نزدیک ترین نقاط داده به خط تصمیم گیری گفته میشود.



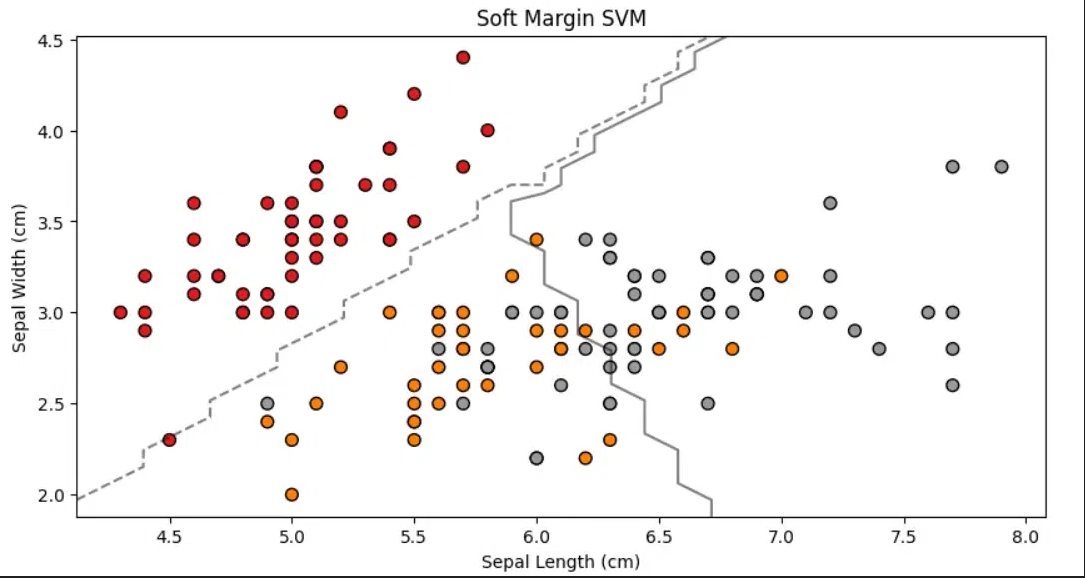
ب)

درست است. تابع kernel در واقع نزدیکی بین 2 داده ورودی را محاسبه میکند. این کار با محاسبه ضرب داخلی دو بردار (vector) در فضای ویژگی (feature space) انجام میشود. فضای ویژگی همان فضای مسئله ما هست اما در بعد های بالاتر یعنی بعد بیشتری به آن اضافه کردیم و مثلا از 2 بعد آن را به 3 بعد بردیم.

**سوال 4.**

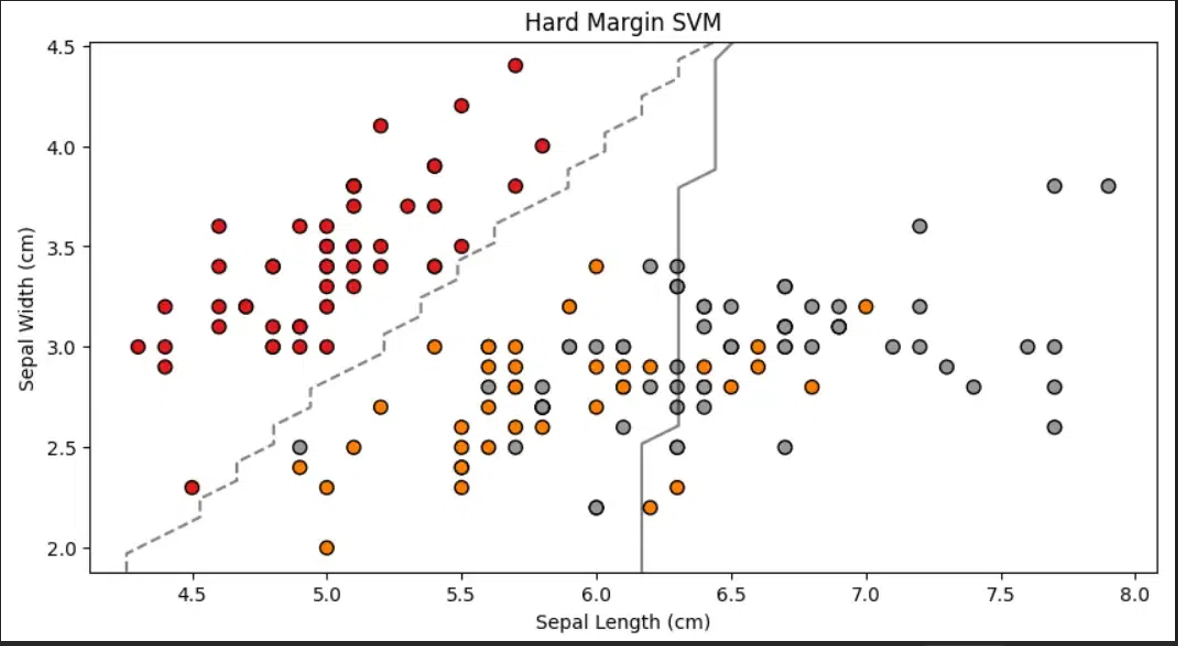
* Soft margin:

یک رویکردی است که اجازه میدهد که misclassification یا خطا در مجموعه آموزشی صورت بگیرد هنگامی که به دنبال مرز تصمیم گیری هستیم. این کار اجازه میدهد تا مدل ما overfit نباشد همچنین مرز تصمیم گیری ما نا پایدار نباشد. با اینکار مدل ما تعمیم پذیر خواهد بود نسبت به داده های جدیدی که آنها را ملاقات نکرده است. هم برای داده های خطی جدایی پذیر و غیر خطی جدا پذیر مناسب هستند. همچنین باعث یک trade offیی بین پیدا کردن بیشترین حد margin و کمترین مقدار خطا صورت بگیرد که اصطلاحا low variance – high bias گفته میشود.



* Hard margin:

در طرف مقابل نسبت به soft، این یک رویکرد سخت گیر هست. که اجازه misclassification نمیدهد. به این معنی که داده ها باید به صورت خطی جدایی پذیر باشند. این رویکرد فقط به دنبال بیشینه کردن margin است و فقط از بردار های پشتیبان کمک میگیرد. در واقع این رویکرد هیچ تحمل پذیری در رابطه با noise ندارد و مستعد overfit شدن است.



**سوال 5.**

در رویکرد آموزش، SVM ها به دنبال پیدا کردن یک مرز تصمیم گیری یا یک صفحه (hyperplane) برای جدا کردن داده های کلاس های مختلف هست که هدف آن بدست آوردن بیشترین margin است. در طرف دیگر شبکه های عصبی با کمک لایه های مختلف و تغییر مداوم پارامترهای آزاد مثل وزن ها یاد میگیرند.

SVM برای استفاده از datasetهای کوچک تا متوسط مناسب هستند مخصوصا برای کلاس بندی های دودویی (binary). در طرف دیگر شبکه های عصبی بسیار نسبت منعطف هستند و میتوانند مدل های پیچیده تری به شمار بیان، و میتوانند روی مجموعه داده های بسیار بزرگ و بعد های بالاتر مورد استفاده قرار بگیرند. به همین دلیل SVM نیاز به منابع محاسباتی کمتری نسبت به شبکه های عصبی دارند. SVM ها هایپر پارامتر های کمتری نسبت به شبکه های عصبی دارند که فقط شامل C و پارامتر های kernel هست. همچنین انتخاب برای kernel هم بین مدل خطی و RBF محدود هستند. از طرف دیگر شبکه های عصبی با طیف گسترده ای از هایپر پارامترها و ... درگیر هستند.

SVM ها برای کلاس بندی متن، کلاس بندی دودویی و مجموعه داده های کوچک تا متوسط مناسب هستند.

در طرف دیگر شبکه های عصبی برای تشخیص الگو های پیچیده، مجموعه داده های بدون ساختار، کلاس بندی چند کلاسه مناسب هستند.

معمولا از SVM در پروژه های مانند spam email detection مورد استفاده قرار میگیرند اما در طرف دیگر شبکه های عصبی برای داده و پروژه های کار با عکس و ... مناسب هستند.