

جزوه دوره (نسخه پژوهشی)

جامپ یادگیری ماشین (Quera College)

نویسنده: نام شما

تاریخ: ۴ اسفند ۱۴۰۴

فهرست مطالب

پیشگفتار

ج

۱ مقدمه

۱

۱	۱.۱	سلام!
۲	۲.۱	قالب کالج
۲	۳.۱	سیستم امتیازات و دریافت گواهی اصلی
۲	۴.۱	یادگیری ماشین چیست؟
۲	۵.۱	چرا پایتون؟
۲	۶.۱	آماده سازی محیط کار
۲	۷.۱	اجرای نوت بوک ها
۳	۸.۱	گوگل کولب
۳	۹.۱	معرفی مجموعه داده

۲ مدیریت پروژه

۴

۴	۱.۲	اهداف فصل
۴	۲.۲	چرخه پروژه
۴	۳.۲	اولویت بندی
۴	۴.۲	سازماندهی تیم
۴	۵.۲	چرا پروژه ها شکست می خورند؟

۳ آماده سازی داده

۵

۵	۱.۳	اهداف فصل
۵	۲.۳	سوالاتی درباره داده
۵	۳.۳	چالش های داده
۵	۴.۳	ویژگی های مجموعه داده ی خوب
۵	۵.۳	تقسیم بندی مجموعه داده
۶	۶.۳	داده های پرت
۶	۷.۳	مقادیر گم شده
۶	۸.۳	مجموعه داده نامتوازن

۷	۴ مهندسی ویژگی
۷	۱.۴ اهمیت
۷	۲.۴ مشخصات ویژگی خوب
۷	۳.۴ ویژگی‌های دسته‌ای
۷	۴.۴ مقادیر گم‌شده در ویژگی‌های دسته‌ای
۷	۵.۴ ویژگی‌های تقویمی
۸	۶.۴ سنتز ویژگی
۸	۷.۴ تغییر مقیاس ویژگی
۸	۸.۴ نشت داده
۸	۹.۴ فوت و فن‌های مهندسی ویژگی
۸	۱۰.۴ کاهش ابعاد
۸	۱۱.۴ انتخاب ویژگی
۸	۱۲.۴ خط لوله
۹	۵ رگرسیون
۹	۱.۵ اهداف فصل
۹	۲.۵ مقدمه
۹	۳.۵ مدل چیست؟
۹	۴.۵ تخمین، تابع هزینه و بهینه‌سازی
۹	۵.۵ رگرسیون خطی
۱۰	۶.۵ ارزیابی
۱۰	۷.۵ رگرسیون چندجمله‌ای
۱۰	۸.۵ عمومیت
۱۰	۹.۵ رگولاریزیشن
۱۱	۶ دسته‌بندی
۱۱	۱.۶ مقدمه
۱۱	۲.۶ رگرسیون لجستیک
۱۱	۳.۶ ارزیابی - قسمت اول
۱۱	۴.۶ ارزیابی - قسمت دوم
۱۱	۵.۶ کراس ولیدیشن
۱۲	۶.۶ نزدیک‌ترین-k همسایه
۱۲	۷.۶ بیز ساده‌لوحانه
۱۲	۸.۶ ماشین بردار پشتیبان
۱۲	۹.۶ هایپراامترها
۱۲	۱۰.۶ آشنایی با کتابخانه‌ی OYH
۱۲	۱۱.۶ درخت تصمیم
۱۲	۱۲.۶ فوت و فن درخت تصمیم
۱۳	۱۳.۶ بیش‌برازش درخت تصمیم

۷	یادگیری تجمعی	۱۴
۱.۷	اهداف فصل	۱۴
۲.۷	مقدمه	۱۴
۳.۷	جنگل تصادفی	۱۴
۴.۷	الگوریتم AdaBoost	۱۴
۵.۷	الگوریتم Boosting Gradient	۱۴
۶.۷	الگوریتم XGboost	۱۵
۷.۷	روش Stacking	۱۵
۸	پروژه اول	۱۶
۱.۸	مقدمه	۱۶
۲.۸	یادداشت‌ها و راه‌حل	۱۶
۹	شبکه عصبی	۱۷
۱.۹	اهداف فصل	۱۷
۲.۹	پرسپترون	۱۷
۳.۹	آموزش پرسپترون	۱۷
۴.۹	پرسپترون چندلایه	۱۷
۵.۹	عمومیت	۱۷
۱۰	یادگیری نظارت‌نشده	۱۸
۱.۱۰	مقدمه	۱۸
۲.۱۰	الگوریتم PCA	۱۸
۳.۱۰	الگوریتم t-SNE	۱۸
۴.۱۰	خوشه‌بندی با k-means	۱۸
۵.۱۰	خوشه‌بندی با k-modes	۱۸
۶.۱۰	خوشه‌بندی با k-prototype	۱۹
۱۱	پروژه دوم	۲۰
۱.۱۱	اهداف فصل	۲۰
۲.۱۱	تعبیه‌ی متن	۲۰
۳.۱۱	فاصله‌ی ویرایش	۲۰
۴.۱۱	معیار شباهت RBO	۲۰
۱۲	بیشتر بدانید	۲۱
۱.۱۲	نمونه‌کاهی با NearMiss	۲۱
۲.۱۲	نمونه‌افزایی با SMOTE	۲۱
۳.۱۲	درخت رگرسیون	۲۱
۲۲	واژه‌نامه	۲۲

فهرست تصاویر

فهرست جداول

پیشگفتار

(.)

فصل ۱

مقدمه

۱.۱ سلام!

سلام؛ ورود شما را به کالج «یادگیری ماشین ۲ | جامپ تکنیکال» خوش آمد می‌گوییم. از اینکه در کوئرا کالج افتخار میزبانی شما را داریم، به خود می‌بالیم

«یادگیری ماشین ۲ | جامپ تکنیکال» کالج سوم از مسیر علم داده و یادگیری ماشین کوئراست که پس از «یادگیری ماشین ۰ | دروازه ورود» و «یادگیری ماشین ۱ | تحلیل داده با پایتون» طراحی شده است.

هدف ما از تدوین کالج این است که شما را به شکل اصولی و گام‌به‌گام با الگوریتم‌های یادگیری ماشین آشنا کنیم؛ به‌طوری که در نهایت بتوانید الگوریتم‌ها را تحلیل کنید، نقاط ضعف و قوت آن‌ها را بشناسید و تشخیص دهید چگونه از آن‌ها برای حل مسائل دنیای واقعی کمک بگیرید! علاوه بر این که با الگوریتم‌های یادگیری ماشین آشنا می‌شوید، تکنیک‌هایی برای استفاده عملی از آن‌ها را نیز خواهید آموخت؛ به عبارت بهتر، «یادگیری ماشین ۲ | جامپ تکنیکال» آمیخته‌ای موزون از دانش علمی و مهارت عملی یادگیری ماشین کلاسیک است!

این کالج مناسب افرادی است که در حد متوسط با پایتون و کتابخانه‌های Numpy و Pandas آشنایی داشته باشند. داشتن دانش حداقلی از آمار احتمال و حسابان (مباحث مشتق) و جبر خطی (ماتریس‌ها) به شما در فراگیری محتوای این کالج کمک می‌کند. به صورت کلی می‌توان گفت اگر با ریاضیات در حد مقطع دبیرستان آشنا باشید، به راحتی می‌توانید به ماجراجویی در این کالج بپردازید! اگر با پیشنیازهای کالج آشنا نیستید، می‌توانید با گذراندن سایر دوره‌های کوئرا کالج، مهارت‌های لازم را کسب کنید.

شما با داشتن پیش‌نیازهای کالج شروع می‌کنید و در انتها پس از مطالعه‌ی درسنامه‌ها و حل تمرین‌ها، به «دانشمند داده» و «مهندس یادگیری ماشین» تبدیل می‌شوید که قطعاً راه طولانی اما شیرینی برای حرفه‌ای شدن پیش رو دارید!

البته برای حفظ تناسب کالج و افزایش احتمال یادگیری عمیق، از آموزش کامل یادگیری عمیق، که زیرمجموعه یادگیری ماشین است، پرهیز شده است؛ بلکه شبکه‌های عصبی مصنوعی و مقدمات یادگیری عمیق را در فصل «شبکه عصبی» آورده‌ایم. برای فراگیری یادگیری عمیق، می‌توانید منتظر کالج‌های بعدی مسیر علم داده و یادگیری ماشین کوئرا کالج باشید

امیدواریم با ارائه‌ی آموزش باکیفیت، گامی مثبت در افزایش دانش و پیشرفت شما برداریم. سختی‌های مسیر نه تنها ناامیدمان نمی‌کند بلکه توان‌مان را بیشتر و تصمیم‌مان را راسخ‌تر می‌کند. مسیر رسیدن به قله‌ای که آرزویش را داریم؛ جوانانی توانمند، پرتلاش و ایرانی پیشرفته...



۲.۱ قالب کالج

[]

۳.۱ سیستم امتیازات و دریافت گواهی اصلی

[]

۴.۱ یادگیری ماشین چیست؟

[]

۵.۱ چرا پایتون؟

[]

۶.۱ آماده‌سازی محیط کار

[]

۷.۱ اجرای نوت‌بوک‌ها

[]

۸.۱ گوگل کولب

[]

۹.۱ معرفی مجموعه داده

[]

فصل ۲

مدیریت پروژه

۱.۲ اهداف فصل

[]

۲.۲ چرخه پروژه

[]

۳.۲ اولویت‌بندی

[]

۴.۲ سازماندهی تیم

[]

۵.۲ چرا پروژه‌ها شکست می‌خورند؟

[]

فصل ۳

آماده‌سازی داده

۱.۳ اهداف فصل

[]

۲.۳ سوالاتی درباره داده

[]

۳.۳ چالش‌های داده

[]

۴.۳ ویژگی‌های مجموعه داده‌ی خوب

[]

۵.۳ تقسیم‌بندی مجموعه داده

[]

۶.۳ داده‌های پرت

]	[
---	---

۷.۳ مقادیر گم‌شده

]	[
---	---

۸.۳ مجموعه داده نامتوازن

]	[
---	---

فصل ۴

مهندسی ویژگی

۱.۴ اهمیت

[]

۲.۴ مشخصات ویژگی خوب

[]

۳.۴ ویژگی‌های دسته‌ای

[]

۴.۴ مقادیر گم‌شده در ویژگی‌های دسته‌ای

[]

۵.۴ ویژگی‌های تقویمی

[]

۶.۴ سنتز ویژگی

[]

۷.۴ تغییر مقیاس ویژگی

[]

۸.۴ نشت داده

[]

۹.۴ فوت و فن های مهندسی ویژگی

[]

۱۰.۴ کاهش ابعاد

[]

۱۱.۴ انتخاب ویژگی

[]

۱۲.۴ خط لوله

[]

فصل ۵

رگرسیون

۱.۵ اهداف فصل

[]

۲.۵ مقدمه

[]

۳.۵ مدل چیست؟

[]

۴.۵ تخمین، تابع هزینه و بهینه‌سازی

[]

۵.۵ رگرسیون خطی

[]

۶.۵ ارزیابی

[]

۷.۵ رگرسیون چند جمله‌ای

[]

۸.۵ عمومیت

[]

۹.۵ رگولاریزیشن

[]

فصل ۶

دسته‌بندی

۱.۶ مقدمه

[]

۲.۶ رگرسیون لجستیک

[]

۳.۶ ارزیابی - قسمت اول

[]

۴.۶ ارزیابی - قسمت دوم

[]

۵.۶ کراس ولیدیشن

[]

۶.۶ نزدیک‌ترین-k همسایه

[]

۷.۶ بیز ساده‌لوحانه

[]

۸.۶ ماشین بردار پشتیبان

[]

۹.۶ هایپراامتراها

[]

۱۰.۶ آشنایی با کتابخانه‌ی OYH

[]

۱۱.۶ درخت تصمیم

[]

۱۲.۶ فوت و فن درخت تصمیم

[]

۱۳.۶ بیش‌برازش درخت تصمیم

[]

فصل ۷

یادگیری تجمعی

۱.۷ اهداف فصل

[]

۲.۷ مقدمه

[]

۳.۷ جنگل تصادفی

[]

۴.۷ الگوریتم AdaBoost

[]

۵.۷ الگوریتم Boosting Gradient

[]

۶.۷ الگوریتم XGboost

[]

۷.۷ روش Stacking

[]

فصل ۸

پروژه اول

۱.۸ مقدمه

[]

۲.۸ یادداشت‌ها و راه‌حل

[]

فصل ۹

شبکه عصبی

۱.۹ اهداف فصل

[]

۲.۹ پرسپترون

[]

۳.۹ آموزش پرسپترون

[]

۴.۹ پرسپترون چندلایه

[]

۵.۹ عمومیت

[]

فصل ۱۰

یادگیری نظارت نشده

۱.۱۰ مقدمه

[]

۲.۱۰ الگوریتم PCA

[]

۳.۱۰ الگوریتم t-SNE

[]

۴.۱۰ خوشه بندی با k-means

[]

۵.۱۰ خوشه بندی با k-modes

[]

۶.۱۰ خوشه‌بندی با k-prototype

[]

فصل ۱۱

پروژه دوم

۱.۱۱ اهداف فصل

[]

۲.۱۱ تعبیه‌ی متن

[]

۳.۱۱ فاصله‌ی ویرایش

[]

۴.۱۱ معیار شباهت RBO

[]

فصل ۱۲

بیشتر بدانید

۱.۱۲ نمونه‌گاهی با NearMiss

[]

۲.۱۲ نمونه‌افزایی با SMOTE

[]

۳.۱۲ درخت رگرسیون

[]

واژه‌نامه

() .

منابع

() . / / /