

به نام خدا

پروژه درس هوش مصنوعی – نیمسال دوم ۱۴۰۰-۱۴۰۱

دانشکده مهندسی صنایع

دانشگاه علم و صنعت ایران

در پروژه پایان ترم درس هوش مصنوعی، انتظار می‌رود دانشجویان یکی از مسایل زیر را بر حسب علاقه انتخاب کرده و با استفاده از آموخته‌های خود در طول ترم، گزارش تحلیلی از چگونگی پیاده‌سازی هوش مصنوعی که وظیفه خواسته شده را انجام می‌دهد، ارائه دهند. لازم است در گزارش به موارد ذیل اشاره شود:

۱. پیش پردازش داده
۲. انتخاب ویژگی‌ها
۳. انتخاب مدل مناسب
۴. تنظیم پارامترهای مدل انتخابی
۵. پیاده سازی مدل نهایی و گزارش نتایج

گزارش نهایی تنها شامل موارد فوق نشده و لازم است تمامی مراحل تحلیل و پیاده‌سازی مدل نهایی در گزارش گنجانده شود. کدهای پیاده‌سازی (ترجیحا پایتون) نیز در کنار توضیحات آورده شده و عملکرد هر یک از دستورات نیز تشریح گردد. در تدوین گزارش لازم است تمامی معیارهای تدوین یک گزارش علمی مانند تنظیم فهرست مطالب، شماره‌گذاری عناوین، شکل‌ها و جداول، استفاده از پانویس و تعاریف واژگان لاتین، مرجع‌دهی استاندارد و رعایت شود. همچنین وجود بخش‌های چکیده و نتیجه‌گیری در گزارش الزامی است. موارد ذیل در ارزیابی پروژه ها مدنظر قرار خواهد گرفت:

موارد ذیل در ارزیابی پروژه ها مدنظر قرار خواهد گرفت:

۱. **ساختار گزارش نهایی:** در نگارش گزارش لازم است تمامی معیارهای تدوین یک گزارش علمی رعایت شود (مشابه مطالب بررسی شده در درس روش تحقیق).
۲. **بررسی کدهای پیاده‌سازی شده:** سیستم هوش مصنوعی پیاده‌سازی شده باید بر روی هر رایانه دیگری نیز بدون خطا ایجاد شود و نتایج مشابه با نتایج گزارش شده را بازتولید کند. به این منظور لازم است تمامی مراحل پیاده‌سازی به صورت یکپارچه در فایل مربوطه آورده شود. توجه داشته باشید بررسی گزارش نهایی منوط به اجرای بدون خطای کد شما خواهد بود. همچنین در صورت استفاده از زبان پایتون با استفاده از کتابخانه pickle مدل های آموزش دیده را ذخیره کرده و به ضمیمه گزارش ارسال فرمایید (ارسال مدل ذخیره شده با استفاده از زبان R و نرم افزار MATLAB هم امکان پذیر است).
۳. **بررسی صحت خروجی ها گزارش شده:** هر کدام از مسایل دارای داده‌های آزمون می‌باشد که در اختیار دانشجویان قرار نگرفته و مدل پیاده‌سازی شده (فایل ذخیره شده توسط pickle) به وسیله آن مورد ارزیابی قرار خواهد گرفت. شاخص دقت به همراه ماتریس بهم‌ریختگی داده‌ها آزمون معیار ارزیابی مدل پیاده‌سازی شده خواهد بود. لازم به ذکر است برای مساله دوم (که از جنس خوشه بندی بوده) خروجی‌های یک مدل پایه (که عملکرد آن قبل صحت سنجی شده) به عنوان برچسب‌های صحیح مدنظر قرار خواهد گرفت.

تذکر مهم: موضوعات پیشنهادی از نظر میزان پیچیدگی به ترتیب بیان شده‌اند. به این ترتیب که مساله اول بیشترین پیچیدگی و مساله آخر کمترین پیچیدگی را خواهد داشت. لذا هنگام ارزیابی پروژه، سختگیری بیشتری برای موضوعات دارای پیچیدگی کمتر لحاظ خواهد شد. موضوعات مدنظر به شرح زیر است:

موضوع اول: دسته‌بندی تصاویر

در این مساله لازم است یک سیستم دارای هوش مصنوعی پیاده‌سازی شود که بتواند تمایز بین تصاویر گربه/سگ را تشخیص داده و برچسب مناسب را برای تصویر جدید تخصیص دهد. تصاویر مورد نیاز به منظور آموزش الگوریتم انتخابی در اختیار دانشجویان قرار خواهد گرفت. این مساله با توجه به مشخص بودن برچسب داده‌های آموزش، یک مساله دسته‌بندی خواهد بود. استفاده از تمامی الگوریتم‌ها به منظور انجام وظیفه مورد نظر بلا مانع است. توجه داشته باشید که علاوه بر موجودیت مورد نظر (سگ و گربه) موجودیت‌های دیگری نیز در تصاویر حضور دارند.

موضوع دوم: پیش‌بینی فروش – Deep Learning

در این مساله دانشجویان باید میزان فروش ماه آتی شرکت را به ازای متغیرهای تقویمی با استفاده از الگوریتم LSTM پیش‌بینی کنند. متغیرهای مستقل در این مساله روز ماه، روز هفته، ماه سال، هفته ماه، هفته سال و روز سال هستند. پیش‌بینی باید به صورت روزانه انجام شده و خروجی‌ها گزارش شوند. برای آموزش شبکه عصبی لازم است از رویکرد یادگیری Cross-validation استفاده شود.

موضوع سوم: پردازش متن – Topic Modeling

در این مساله دانشجویان باید یک سیستم هوش مصنوعی پیاده‌سازی کنند که بتواند دسته‌بندی تیترهای اصلی اخبار را انجام دهد. در این مساله متن مربوط به اخبار یک شبکه به همراه تاریخ نشر آن در اختیار دانشجویان قرار خواهد گرفت. داده‌های این مساله دارای برچسب نیستند و لذا ابتدا لازم است عناوین موجود خوشه‌بندی شده و سپس برچسب مناسب هر خوشه با توجه به مفاد آن تعیین شود. توجه داشته باشید که هر یک از خوشه‌ها می‌تواند چندین برچسب به خود اختصاص دهد. در نهایت، سیستم پیاده‌شده باید عنوان جدید را به صورت یک رکورد جدید مشابه با داده‌های ورودی دریافت کرده و برچسب آن را مشخص کند.

موضوع چهارم: پردازش متن – Text Classification

در این مساله دانشجویان باید یک سیستم هوش مصنوعی را طراحی و آموزش دهند که بتواند برچسب Spam یا Not-Spam را به متن پیامک دریافتی تخصیص دهد. داده مورد نیاز برای آموزش هوش مصنوعی ۵۷۰۰ پیامک به زبان انگلیسی می‌باشد که برچسب هر یک پیامک‌ها به آن‌ها اختصاص داده شده است. در این مساله لازم است ابتدا پیش پردازش متن صورت گرفته و سپس با استخراج ویژگی سیستم هوش مصنوعی برچسب گذاری را انجام دهد.

موفق باشید