چرخه زندگی یادگیری ماشین

what is ML lifecycle?

machine learning life cycle is defined as a cyclical process which involves three-phase, process data, training phase, and inference phase acquired by the data scientist and the data engineers to develop,

train and serve the models using the huge amount of data that are involved in various applications

**چرخه عمر ML چیست؟**

چرخه زندگی یادگیری ماشین، به عنوان یک فرایند چرخه‌ای تعریف می شود که شامل سه فاز، داده‌های فرایند، مرحله آموزش و مرحله استنتاج می‌شود که توسط دانشمندان‌ و مهندسانِ داده برای توسعه ایجاد می‌شود. با استفاده از حجم عظیمی از داده‌ها که در برنامه‌های مختلف دخیل هستند، مدل‌ها را آموزش داده و سرویس می‌دهند.

steps involved in ML lifecycle

define project objectives

gathering data

data preparation

model training

model testing

deploy models

model inference

monitor and optimize

**مراحل مربوط به چرخه زندگی ML**

۱) تعیین اهداف پروژه

۲) جمع آوری اطلاعات

۳) آماده سازی داده‌ها

۴) آموزش مدل‌ها

۵) تست مدل‌ها

۶) استقرار مدل‌ها

۷) استنباط مدل‌ها

۸) نظارت و بهینه‌سازی

define the problem

the first step of the life cycle is to understand the problem and to know the purpose of the problem.

therefore, before starting the life cycle, we need to understand the problem beacuse the good result depends on the better underestanding of the problem.

**مشکل را تعریف کنید**

اولین مرحله از چرخه زندگی درک مسئله و دانستن هدف مشکل است.

بنابراین، قبل از شروع چرخه زندگی، باید مسئله را درک کنیم زیرا هرچه مسئله بهتر درک شود، نتیجه بهتری به‌دست می‌آید.

gathering data

the next step is to identify, collect and prepaare all of the relevant data for use in machine learning.

in this step, we need to identify the different data sources, as data can be collected from various sources such as files.

database, internet, or mobile devices.the quantity and quality of the collected data will determine the effciency of the output.

**جمع آوری اطلاعات**

گام بعدی، شناسایی، جمع آوری و آماده‌سازی تمام داده‌های مربوطه برای استفاده در یادگیری ماشین است.

در این مرحله ما باید منابع مختلف داده را شناسایی کنیم، زیرا داده‌ها می‌توانند از منابع مختلف مانند پرونده‌ها جمع آوری شوند.

پایگاه داده ، اینترنت یا دستگاه های تلفن همراه. کمیت و کیفیت داده‌های جمع آوری شده اثر خروجی را تعیین می کند.

data preparation

make sure your data is clean, secure, and governed. it is the process of cleaning the data, selecting the variable to use, and transforming the data in a proper format to make it more suitable for analysis in the next step.

you can also do feature engineering or feature selection which helps to identify the important features within a dataset.

**آماده سازی داده‌ها**

اطمینان حاصل کنید که داده‌های شما مرتب، ایمن و مدیریت‌شده هستند. این فرآیند مرتب‌کردن داده‌ها، انتخاب متغیر مورد استفاده و تبدیل داده‌ها در یک قالب درست، در گام‌های بعدی تجزیه‌ و تحلیل راحت‌تری را به ما می‌دهد.

است تا در مرحله بعدی برای تجزیه و تحلیل مناسب‌تر باشد.

شما همچنین می توانید مهندسی ویژگی یا انتخاب ویژگی را انجام دهید که به شناسایی ویژگی های مهم در یک مجموعه داده کمک می کند.

model training

we need to select the models to try and the selection depends on the business problem we are handling or more that depends on the application and end results.

we also do hyper-parameter tuning.

tuning of model parameter depends on multiple aspects like cross-validation, outliter or noisy data removal etc.

**آموزش مدل**

ما برای مشکل تجاری خود باید مدل‌ها را انتخاب کنیم، که این انتخاب به برنامه و نتایج نهایی بستگی دارد.

ما همچنین تنظیم بیش از حد پارامتر را انجام می دهیم.

تنظیم پارامتر مدل به جنبه‌های مختلفی از جمله اعتبارسنجی متقابل‌، حذف داده‌ها یا حذف داده‌های پر سر و صدا و غیره بستگی دارد.

Model testing

the developed model has to be tested on the unseen data before deployed into the field or production environments.there are various KPIs available in the machine learning area for a model which can vary on the basis of models.

**تست مدل**

مدل توسعه یافته باید قبل از استقرار در محیط‌های تولید بر روی داده‌های دیده‌نشده آزمایش شود.

KPI های مختلفی در منطقه یادگیری ماشین برای یک مدل وجود دارد که می تواند براساس مدل‌ها متفاوت باشد.

model deployment

trained model has to be pickled before the deployment which is a platform independent executable in layman terms.

the pickled model object can be deployed using various methods ike rest APIs or Micro-services.

**استقرار مدل**

مدل آموزش دیده باید قبل از استقرار ترشی شود که از نظر عرفی یک پلت فرم قابل اجرا است.

آبجکت مدل ترشی را می توان با استفاده از روشهای مختلف مانند API های استراحت یا خدمات خرد استفاده کرد.

monitor and optimize

once a model is deployed, there are a number of measures that can be taken to improve robustness and learning project to be successful in the long term , it requires more attention with regards to lineage, monitoring, testing and model drift. these key components are often lacking due to missing tooling, inexperience and relatively high development costs.

**نظارت و بهینه سازی**

هنگامی که یک مدل به کار گرفته شد‌، تعدادی از اقدامات برای بهبود قدرت و یادگیری وجود دارد تا پروژه در بلند مدت موفقیت آمیز باشد، این امر به توجه بیشترنسب به نظارت‌، آزمایش و رانش مدل نیاز دارد. این مولفه‌های کلیدی اغلب به دلیل از دست دادن ابزار‌، بی‌تجربگی و هزینه‌های نسبتاً بالای توسعه‌، فاقد آن هستند.