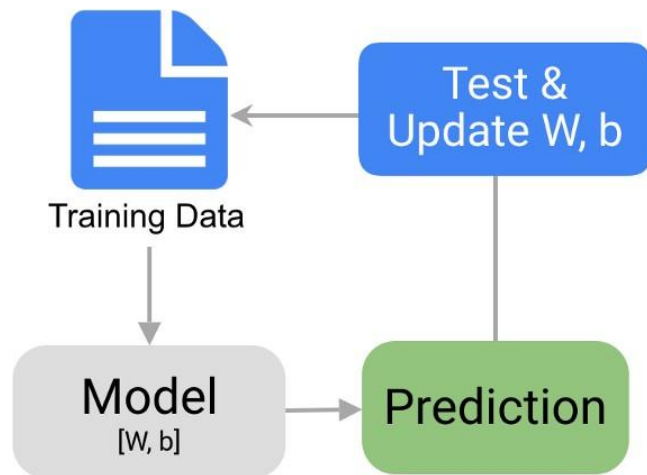


### Maşın Öyrənməsinin Tətbiq Addımları



#### Maşın öyrənmə Həyat dövrü

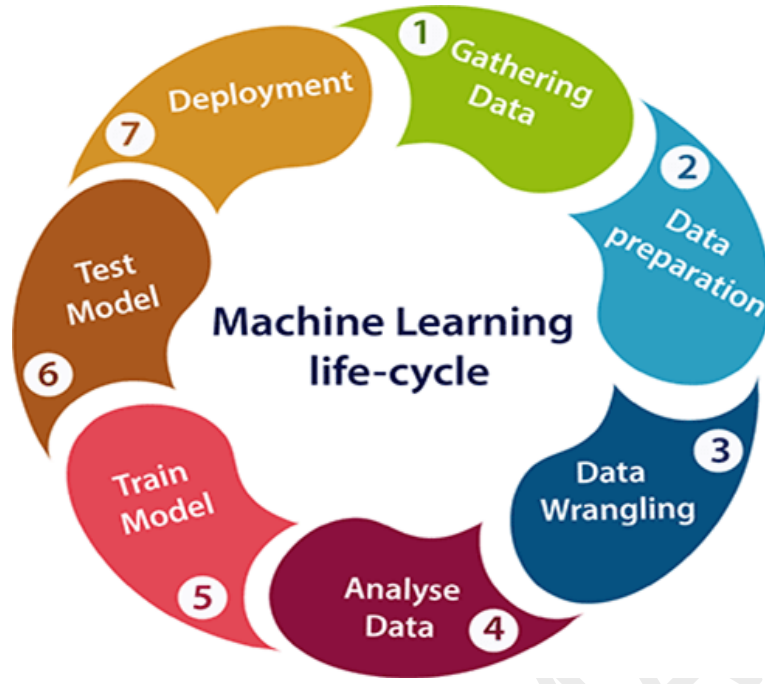
Maşın öyrənməsi kompüter sistemlərinə açıq şəkildə proqramlaşdırılmadan avtomatik öyrənmək imkanı verir. Bəs maşın öyrənmə sistemi necə işləyir? Buna görə də, maşın öyrənməsinin həyat dövrü istifadə edərək təsvir edilə bilər. Maşın öyrənməsinin həyat dövrü səmərəli maşın öyrənməsi layihəsi yaratmaq üçün tsiklik bir prosesdir. Həyat dövrünün əsas məqsədi problemin və ya layihənin həllini tapmaqdır.

#### Maşın öyrənməsinin həyat dövrü aşağıda verilmiş yeddi əsas addımdan ibarətdir:

- Dataların toplanması
- Dataların hazırlanması
- Dataların müzakirəsi
- Dataların təhlil edilməsi
- Dataların train edilməsi
- Dataların test edilməsi
- Tətbiq

Bütün prosesdə ən vacib şey problemi başa düşmək və problemin məqsədini bilməkdir. Buna görə də, həyat dövrünə başlamazdan əvvəl problemi başa düşməliyik, çünki yaxşı nəticə problemin daha yaxşı başa düşülməsindən asılıdır.

Həyat dövrü ərzində bir problemi həll etmək üçün "model" adlı maşın öyrənmə sistemi yaradıırıq və bu model "train" verilərək yaradılır. Ancaq modeli hazırlamaq üçün bizə data lazımdır, ona görə də onun həyat dövrü məlumatların toplanması ilə başlayır.



**Zəkanın maşınlara ötürülməsi vəzifəsini 7 əsas mərhələyə bölmək olar:**

#### **1. Data toplaması:**

Bildiyiniz kimi, Maşınlar əvvəlcə verdiyiniz məlumatlardan öyrənirlər. Maşın öyrənmə modelinizin düzgün nümunələri tapması üçün etibarlı dataların toplanması son dərəcə vacibdir. Maşına verdiyiniz məlumatların keyfiyyəti modelinizin nə qədər dəqiq olduğunu müəyyənləşdirəcəkdir. Yanlış və ya köhnəlmiş datalarınız varsa, əlaqəli olmayan səhv nəticələr və ya proqnozlarınız olacaq.

Etibarlı mənbədən alınan datalardan istifadə etdiyinizə əmin olun, çünki bu, modelinizin nəticələrinə birbaşa təsir edəcəkdir. Yaxşı data müvafiqdir, bir neçə əskik və təkrar dəyərə malikdir və mövcud müxtəlif alt kateqoriyaların/siniflərin yaxşı təmsilidir.

Dataların toplanması maşın öyrənməsinin həyat dövrünün ilk addımıdır. Bu addımın məqsədi məlumatlarla bağlı hər hansı problemi müəyyən etmək və ələ keçirməkdir. Bu addımda biz müxtəlif məlumat mənbələrini müəyyən etməliyik, çünki məlumatlar fayllar, verilənlər bazası, internet və ya mobil cihazlar kimi müxtəlif mənbələrdən toplana bilər. Bu, həyat dövrünün ən vacib addımlarından biridir. Toplanan dataların kəmiyyəti və keyfiyyəti çıxışın səmərəliliyini müəyyən edəcəkdir. Nə qədər çox məlumat varsa, proqnoz bir o qədər dəqiqdir. Bu addıma aşağıdakı tapşırıqlar daxildir:

- Müxtəlif data mənbələrini müəyyən edin
- Data toplayın
- Müxtəlif mənbələrdən dataları inteqrasiya edin

Yuxarıdakı tapşırığı yerinə yetirməklə, biz verilənlər set adlanan ardıcıl məlumat seti əldə edirik. Sonrakı addımlarda istifadə olunacaq.

### 2. Dataların hazırlanması:

Dataları topladıqdan sonra onu növbəti addımlara hazırlamalıyıq. Dataların hazırlanması məlumatlarımızı uyğun yerə yerləşdirdiyimiz və maşın öyrənmə təlimimizdə istifadə üçün hazırladığımız bir addımdır. Bu mərhələdə bütün datalar əvvəlcə bir araya gətirilir və sonra verilənlərin sıralanması təsadüfiləşdirilir. Bu addımı iki prosesə bölmək olar: Data kəşfiyyatı: işləməli olduğumuz dataların təbiətini anlamaq üçün istifadə olunur. Biz dataların xüsusiyyətlərini, formatını və keyfiyyətini başa düşməliyik. Dataların daha yaxşı başa düşülməsi effektiv nəticəyə gətirib çıxarır. Bunun içərisində biz korrelyasiyaları, ümumi tendensiyaları və kənar göstəriciləri tapırıq. Datanın əvvəlcədən işlənməsi: İndi növbəti addım verilənlərin təhlil üçün əvvəlcədən emal edilməsidir. Datalarınız olduqdan sonra onu hazırlamalısınız. Bunu belə edə bilərsiniz:

- Əlinizdə olan bütün dataları birləşdirir və təsadüfiləşdirir. Bu, dataların bərabər paylanmasına və sifarişin öyrənmə prosesinə təsir etməməsinə kömək edir.
- Arzuolunmaz dataları, çatışmayan dəyərləri, sətirləri və sütunları, dublikat dəyərləri, məlumat növünün çevrilməsini və s. silmək üçün dataların təmizlənməsi. Siz hətta data dəstini yenidən qurmaq və sətir və sütunları və ya sətir və sütunların indeksini dəyişməli ola bilərsiniz.
- Verilənlərin necə qurulduğunu anlamaq və müxtəlif dəyişənlər və mövcud siniflər arasındakı əlaqəni anlamaq üçün dataları vizuallaşdırın.
- Təmizlənmiş dataları iki yerə bölün - train dəsti və test dəsti. Train dəsti modelinizin öyrəndiyi dəstdir. Train-dən sonra modelinizin düzgünlüyünü yoxlamaq üçün test dəsti istifadə olunur.

Data mübahisəsi xam dataların təmizlənməsi və istifadə edilə bilən formata çevrilməsi prosesidir. Verilənlərin təmizlənməsi, istifadə olunacaq dəyişənin seçilməsi və növbəti mərhələdə təhlil üçün daha uyğun olması üçün məlumatların müvafiq formata çevrilməsi prosesidir. Bu, bütün prosesin ən vacib addımlarından biridir. Keyfiyyət problemlərini həll etmək üçün məlumatların təmizlənməsi lazımdır.

Bəzi datalar faydalı olmaya biləcəyi üçün topladığımız datalar mütləq istifadəmiz üçün deyil. Real dünya tətbiqlərində toplanmış dataların müxtəlif problemləri ola bilər, o cümlədən:

- Çatışmayan Dəyərlər
- Dublikat məlumat
- Yanlış data
- Səs-küylü

## MÜƏLLİF: İBRAHİM İSMAYİL

Buna görə məlumatları təmizləmək üçün müxtəlif filtrləmə üsullarından istifadə edirik. Yuxarıda göstərilən problemləri müəyyən etmək və aradan qaldırmaq vacibdir, çünki onlar nəticənin keyfiyyətinə mənfi təsir göstərə bilər.

**Dataların Təhlili:** İndi təmizlənmiş və hazırlanmış datalar təhlil mərhələsinə keçir. Bu addım daxildir:

- Analitik üsulların seçilməsi
- Bina modelləri
- Nəticəni nəzərdən keçirmək

Bu addımın məqsədi müxtəlif analitik üsullardan istifadə edərək məlumatları təhlil etmək və nəticəni nəzərdən keçirmək üçün maşın öyrənmə modelini yaratmaqdır. Təsnifat, Reqressiya, Klaster analizi, Assosiasiya və s. Bu, maşın öyrənmə üsullarını seçdiyimiz problemlərin növünü müəyyənləşdirməklə başlayır, sonra hazırlanmış məlumatlardan istifadə edərək model yaratmaq və modeli qiymətləndirmək. Buna görə də, bu addımda biz məlumatları götürürük və modeli qurmaq üçün maşın öyrənmə alqoritmlərindən istifadə edirik.

### 3. Model Seçimi:

Alqoritmlər verilənlər toplusunun strukturuna və problemin növünə görə müəyyən edilir. Problemin növündən asılı olaraq hansı maşın öyrənmə alqoritminin seçiləcəyinə qərar vermək üçün təcrübə və bacarıq əldə etməyə ehtiyac var. Maşın öyrənmə modeli toplanmış məlumatlar üzərində maşın öyrənməsi alqoritmini işlətdikdən sonra əldə etdiyiniz nəticəni müəyyən edir. Qarşıya qoyulan vəzifəyə uyğun bir model seçmək vacibdir. Bu illər ərzində alimlər və mühəndislər nitqin tanınması, təsvirin tanınması, proqnozlaşdırılması və s. Müxtəlif vəzifələr üçün uyğun olan müxtəlif modellər hazırladılar, məsələn: Bundan əlavə, modelinizin ədədi və ya kateqoriyalı məlumatlara uyğun olub olmadığını görməli və ona uyğun seçim etməlisiniz.

**Train Modeli:** İndi növbəti addım modeli öyrətməkdir, bu addımda biz modelimizi problemin daha yaxşı həlli üçün onun performansını yaxşılaşdırmaq üçün öyrədirik. Biz müxtəlif maşın öyrənmə alqoritmlərindən istifadə edərək modeli öyrətmək üçün məlumat dəstlərindən istifadə edirik. Modeli öyrətmək onun müxtəlif nümunələri, qaydaları və xüsusiyyətləri anlama bilməsi üçün zəruridir.

**Test Modeli:** Maşın öyrənmə modelimiz xüsusi verilənlər bazasında öyrədildikdən sonra modeli sınaqdan keçiririk. Bu addımda modelimizə test məlumat dəsti təqdim edərək modelimizin düzgünlüyünü yoxlayırıq.

Modelin sınaqdan keçirilməsi layihənin və ya problemin ehtiyacından asılı olaraq modelin faiz dəqiqliyini müəyyən edir.

### 4. Modelin Treyninqi:

Train maşın öyrənməsində ən vacib addımdır. Train-ə nümunələri tapmaq və proqnozlar vermək üçün hazırlanmış məlumatları maşın öyrənmə modelinizə qidalandırırırsınız. Bu, modelin verilənlərdən öyrənilməsi ilə nəticələnir ki, o, tapşırıq dəstini yerinə yetirə bilsin. Zaman keçdikcə, train-lə, model proqnozlaşdırmaqda daha yaxşı olur.

### 5. Modelin qiymətləndirilməsi:

Modelinizi öyrətdikdən sonra onun necə işlədiyini yoxlamaq lazımdır. Bu, modelin performansını əvvəllər görünməmiş məlumatlar üzərində sınaqdan keçirməklə həyata keçirilir. İstifadə olunan görünməyən məlumatlar əvvəllər məlumatlarımızı böldüyünüz test dəstidir. Test train üçün istifadə edilən eyni verilənlər üzərində aparılırsa, siz dəqiq ölçmə əldə etməyəcəksiniz, çünki model artıq verilənlərlə tanış olub və əvvəllər olduğu kimi eyni nümunələri tapıb. Bu sizə qeyri-mütənasib yüksək dəqiqlik verəcəkdir. Test məlumatlarında istifadə edildikdə, modelinizin necə performans göstərəcəyinə və sürətinə dair dəqiq ölçü əldə edirsiniz.

**Yerləşdirmə:** Maşın öyrənməsinin həyat dövrünün son mərhələsi yerləşdirmədir, burada modeli real dünya sisteminə yerləşdiririk.

Hazırlanmış model məqbul sürətlə ehtiyaclarımıza uyğun dəqiq nəticə verirsə, modeli real sistemə yerləşdiririk. Lakin layihəni tətbiq etməzdən əvvəl onun mövcud məlumatlardan istifadə edərək performansını yaxşılaşdırıb-yaxşımadığını yoxlayacağıq. Yerləşdirmə mərhələsi layihə üçün yekun hesabatın hazırlanmasına bənzəyir.

**6. Parametrlərin qurulması:**

Modelinizi qurduqdan və qiymətləndirdikdən sonra dəqiqliyini hər hansı bir şəkildə yaxşılaşdırmağın mümkün olub olmadığını görün. Bu, modelinizdə mövcud olan parametrləri tənzimləməklə edilir. Parametrlər, proqramçının ümumiyyətlə həll etdiyi modeldəki dəyişənlərdir. Parametrinizin müəyyən bir dəyərində dəqiqlik maksimum olacaqdır. Parametr qəbulu bu dəyərlərin tapılmasına aiddir. Optimallaşdırılmalı olan diapozon parametrləri müəyyən edilir.

**7. Proqnozların edilməsi**

Nəhayət, dəqiq proqnozlar vermək üçün modeliniz gözlənilməz, gözlənilən məlumatlar üzərində manipulyasiya edilə bilər.