

AZƏRBAYCAN
RESPUBLİKASININ
DÖVLƏT
STANDARTI

AZS ISO/IEC
23053:2024

Birinci nəşr
2024

Maşın öyrənməsindən istifadə
edən süni intellekt (Sİ)
sistemləri üçün çərçivə sənədi

Framework for Artificial Intelligence
(AI) Systems Using Machine
Learning (ML)



Bu standart Azərbaycan Standartlaşdırma İnstitutunun icazəsi olmadan tam və ya hissə-hissə yenidən çap oluna, çoxaldıla və yayıla bilməz

Elçin İsaqzadə küç., 7-ci köndələn

Qaynar xətt: +994125149308

Email: office@azstand.gov.az

MÜQƏDDİMƏ

1. Bu standart Azərbaycan Elm Fondunun qrant layihəsi (Qrant №AEF-MQM-QA-1-2021-4(41)-8/03/1) çərçivəsində işlənib hazırlanıb və “İnformasiya-kommunikasiya texnologiyaları” standartlaşdırma üzrə Texniki Komitə (AZSTAND/TK 05) tərəfindən təqdim edilib.
2. Azərbaycan Standartlaşdırma İnstitutunun 2024-cü il tarixli sayılı qərarı ilə təsdiq edilib.
3. Bu standart Beynəlxalq Standart ISO/IEC 243053 nəşr 1.0 (2022-06) ilə eynidir.
4. Dövlət standartında müəyyən edilən tələblərin beynəlxalq standartlara, norma, qayda və tövsiyələrə və digər dövlətlərin müvafiq mütərəqqi milli standartlarına, elm, texnika və texnologiyanın müasir nailiyyətlərinə əsaslanmasını müəyyən etmək üçün standartın dövrü yoxlama müddəti ildə 1 dəfədir.

MÜNDƏRİCAT

ÖN SÖZ	6
GİRİŞ	7
1 TƏTBİQ SAHƏSİ	8
2 NORMATİV İSTİNADLAR	8
3 TERMİN VƏ ANLAYIŞLAR	8
3.1 Modelin işlənməsi və istifadəsi	9
3.2 Alətlər	10
3.3 Verilənlər	10
4 QISALTMALAR	11
5 İCMAL	13
6 MAŞIN ÖYRƏNMƏSİ SİSTEMİ	13
6.1 Ümumi müddəalar	13
6.2 Tapşırıq	15
6.2.1 Ümumi müddəalar	15
6.2.2 Reqressiya	15
6.2.3 Klassifikasiya.....	15
6.2.4 Klasterizasiya	16
6.2.5 Anomaliyaların aşkarlanması	16
6.2.6 Ölçülərin azaldılması	16
6.2.7 Digər tapşırıqlar.....	17
6.3 Model	17
6.4 Verilənlər	18
6.5 Alətlər	20
6.5.1 Ümumi müddəalar	20
6.5.2 Verilənlərin hazırlanması	20
6.5.3 Maşın öyrənməsi alqoritmlərinin kateqoriyaları	20
6.5.3.1 Ümumi müddəalar.....	20
6.5.3.2 Neyron şəbəkələri	21
6.5.3.3 Bayes şəbəkəsi.....	25
6.5.3.4 Naïve Bayes alqoritmı	26
6.5.3.5 Dəstək vektoru maşını	26
6.5.3.6 Qərarlar ağacı	26
6.5.4 Maşın öyrənməsinin optimallaşdırma metodları	27
6.5.4.1 Ümumi müddəalar.....	27
6.5.4.2 Qradyent enişi metodları	27
6.5.4.3 Nyuton metodu.....	28
6.5.4.4 Qoşma qradyent.....	28
6.5.4.5 Qauss prosesləri	28
6.5.4.6 Ən kiçik kvadratlar metodu ilə approksimasiya	28
6.5.4.7 Maksimum ehtimalın qiymətləndirilməsi.....	28

6.5.4.8 Gözləmənin maksimallaşdırılması	28
6.5.5 Maşın öyrənməsinin qiymətləndirilməsi metrikaları	29
6.5.5.1 Ümumi müddəalar	29
6.5.5.2 Dəqiqlik, dolğunluq, həssaslıq və spesifiklik	29
6.5.5.3 F1 meyarı	30
6.5.5.4 Doğruluq	30
6.5.5.5 Qəbuledicinin iş xüsusiyyətləri və əyri altındakı sahə	30
6.5.5.6 Xətarlar matrisi	31
6.5.5.7 Kappa əmsalı	32
6.5.5.8 Metyu korrelyasiya əmsalı	32
7 MAŞIN ÖYRƏNMƏSİ YANAŞMALARİ	33
7.1 Ümumi müddəalar	33
7.2 Müəllimlə maşın öyrənməsi	34
7.3 Müəllimsiz maşın öyrənməsi	35
7.4 Qismən müəllimlə maşın öyrənməsi	36
7.5 Öz-özünə maşın öyrənməsi	36
7.6 Möhkəmləndirici maşın öyrənməsi	37
7.7 Transfer öyrənmə	38
8 MAŞIN ÖYRƏNMƏSİ KONVEYERİ	39
8.1 Ümumi müddəalar	39
8.2 Verilənlərin toplanması	41
8.3 Verilənlərin hazırlanması	41
8.4 Modelləşdirmə	44
8.5 Verifikasiya və validasiya	45
8.6 Modelin quraşdırılması	46
8.7 İstismar mərhələsi	47
8.8 ML konveyeri əsasında maşın öyrənməsi prosesinə dair nümunə	47
ƏLAVƏ A	50
ƏDƏBİYYAT	53

ÖN SÖZ

ISO (Beynəlxalq Standartlaşdırma Təşkilatı) və IEC (Beynəlxalq Elektrotexniki Komissiya) dünya üzrə standartlaşdırma sahəsində ixtisaslaşmış sistemi formalaşdırırlar. ISO və ya IEC üzvü olan milli orqanlar texniki fəaliyyətin konkret sahələri ilə məşğul olmaq üçün müvafiq təşkilat tərəfindən yaradılmış texniki komitələr vasitəsilə beynəlxalq standartların hazırlanmasında iştirak edirlər. ISO və IEC texniki komitələri qarşılıqlı maraq doğuran sahələrdə əməkdaşlıq edirlər. ISO və IEC ilə əməkdaşlıq edən digər beynəlxalq təşkilatlar, dövlət və qeyri-hökumət təşkilatları da bu işdə iştirak edirlər.

Bu standartın hazırlanması üçün istifadə olunan və həmçinin sonrakı texniki xidmət üçün nəzərdə tutulan prosedurlar ISO/IEC Direktivlərinin 1-ci hissəsində təsvir edilmişdir. Müxtəlif növ sənədlər üçün tələb olunan fərqli təsdiq meyarlarına xüsusilə diqqət yetirilməlidir. Bu sənəd ISO/IEC Direktivlərinin 2-ci hissəsində (www.iso.org/directives və ya www.iec.ch/members_experts/refdocs) verilmiş qaydalara uyğun olaraq hazırlanmışdır.

Bu standartın bəzi elementlərinin patent hüquqlarının predmeti ola bilməsi diqqət çəkir. ISO və IEC bu patent hüquqlarının hər hansı birinin və ya hamısının müəyyən edilməsinə görə məsuliyyət daşımır. Sənədin hazırlanması zamanı müəyyən edilmiş patent hüquqlarının təfərrüatları “Giriş”də və/və ya ISO və IEC-in (www.iso.org/patents və patents.iec.ch) patent bəyannamələrinin siyahısında təqdim olunur.

Bu standartdakı ticarət adları (“trade name”) haqqında məlumatlar istifadəçilərin rahat istifadəsi üçün təqdim olunur və bu təqdimat tövsiyə xarakteri daşımır.

Standartların könüllü xarakter daşması, uyğunluğun qiymətləndirilməsi üzrə ISO-nun xüsusi termin və ifadələrinin mənası ilə bağlı izahlar, eləcə də Ticarətdə Texniki Maneələrin (Technical Barriers to Trade, TBT) aradan qaldırılması ilə əlaqədar ISO-nun Ümumdünya Ticarət Təşkilatının (ÜTT) prinsiplərinə sadıqlığı haqqında məlumat “www.iso.org/iso/foreword.html” internet informasiya ehtiyatından əldə edilə bilər. IEC ilə bağlı “www.iec.ch/understanding-standards” internet informasiya ehtiyatına müraciət etmək olar.

Bu sənəd ISO/IEC JTC 1 “İnformasiya texnologiyaları” Birgə Texniki Komitəsinin “SC 42, Süni intellekt” Altkomitəsi tərəfindən hazırlanmışdır.

Bu sənədlə bağlı istənilən rəy və suallar milli standartlar üzrə quruma yönəldilməlidir. Bu qurumların tam siyahısı ilə “www.iso.org/members.html” internet informasiya ehtiyatında tanış olmaq olar.

GİRİŞ

Süni intellekt (Sİ) sistemləri, ümumiyyətlə, insan tərəfindən müəyyən edilmiş məqsədlərə nail olmaq üçün kontent, proqnozlar, tövsiyələr və ya qərarlar kimi çıxış verilənləri generasiya edən mühəndis sistemləridir. Sİ bu mürəkkəb məsələlərin həllinə müxtəlif yanaşmaları əks etdirən geniş spektrli texnologiyaları əhatə edir.

Maşın öyrənməsi Sİ-nin sistemlərin verilənlərə və ya təcrübəyə əsaslanaraq öyrənilməsinə imkan yaradan hesablama üsullarından istifadə edən bir sahəsidir. Başqa sözlə, maşın öyrənməsi (ML) sistemləri alqoritmləri təlim verilənlərinə uyğunlaşdırmaq üçün optimallaşdırmaqla və ya “mənfəət” funksiyasının maksimallaşdırılması əsasında alqoritmin məhsuldarlığını artırmaqla işlənilir. Bu standartda maşın öyrənməsi metodlarına aid olan dərin öyrənmədən də bəhs edilir.

Sənəddə bilik, öyrənmə və qərarlar kimi terminlər istifadə olunur. Bununla belə, sənəd maşın öyrənməsinin (ML) antropomorfizasiya edilməsini nəzərdə tutmur.

Bu standart maşın öyrənməsindən istifadə edən Sİ sistemlərini təsvir etmək üçün çərçivə sənədinin hazırlanması məqsədi daşıyır. Standart Sİ sistemləri üçün ümumi terminologiya və ümumi anlayışları hazırlamaqla, sistemlərin aydın təsviri, həmçinin onların işlənməsi və istifadəsi ilə bağlı tətbiq olunan müxtəlif mülahizələr üçün əsasları müəyyən edir. Bu sənəd sahə üzrə ekspertlər və təcrübəsi olmayanlar da daxil olmaqla, geniş auditoriya üçün nəzərdə tutulur. Bununla belə, bəzi bəndlər (5-ci bölmədəki icmalda müəyyən edilən) daha dərin texniki təfərrüatları ehtiva edir.

Bu standart, həmçinin ML sistemlərinin və onların komponentlərinin xüsusi aspektləri ilə bağlı digər standartlar üçün əsasları təmin edir.

AZƏRBAYCAN RESPUBLİKASININ DÖVLƏT STANDARTI

Maşın öyrənməsindən (ML) istifadə edən
süni intellekt (AI) sistemləri üçün çərçivə sənədi

AZE ISO/IEC 23053:2024

Framework for Artificial Intelligence (AI)
Systems Using Machine Learning (ML)

Tətbiq edilmə tarixi ... 2024-cü il

1 TƏTBİQ SAHƏSİ

Bu standart ML texnologiyasından istifadə edən süni intellekt (Sİ) sistemlərinin ümumi şəkildə təsvir edilməsi məqsədilə Sİ və maşın öyrənməsinə (ML) dair çərçivə sənədidir. Sənəd Sİ ekosistemində sistemin komponentlərini və funksiyalarını müəyyən edir. Bu standart Sİ sistemlərini tətbiq və ya istifadə edən dövlət orqanları (qurumları), hüquqi şəxslər də (kommersiya və qeyri-kommersiya təşkilatları) daxil olmaqla, müxtəlif növ və strukturlu təşkilatlara şamil edilir.

2 NORMATİV İSTİNADLAR

Aşağıdakı sənədə mətn boyu istinad onun məzmununun tam və ya qismən bu sənədin tələblərinə uyğun gəldiyi təqdirdə edilmişdir. Tarix qeyd edilmiş istinadlarda yalnız istinad olunan nəşrlər istifadə edilir. Tarix qeyd edilməmiş istinadlarda istinad olunan sənədin (düzəlişlər daxil olmaqla) sonuncu nəşri istifadə edilir.

AZE ISO/IEC 22989:2023, *Informasiya texnologiyaları — Süni intellekt — Süni intellekt üzrə anlayışlar və terminologiya*

3 TERMİN VƏ ANLAYIŞLAR

Bu sənədin məqsədləri üçün AZE ISO/IEC 22989:2023 standartında və aşağıdakı terminlər və anlayışlar istifadə edilir.

ISO və IEC-in standartlaşdırma sahəsində istifadə olunan terminoloji məlumat bazaları aşağıdakı ünvanlarda saxlanılır:

- ISO Onlayn baxış platforması: <https://www.iso.org/obp>;
- IEC Electropedia: <http://www.electropedia.org/>.

3.1 Modelin işlənməsi və istifadəsi

3.1.1

klassifikasiya modeli

<maşın öyrənməsi> kontekstində – Verilmiş giriş üçün gözlənilən çıxışı bir və ya daha çox sinifdən ibarət olan maşın öyrənməsi modeli

3.1.2

regressiya modeli

Verilmiş giriş üçün gözlənilən çıxışı bu giriş verilənlərinin kəsilməz funksiyası olan maşın öyrənməsi modeli

[MƏNBƏ: ISO/IEC TR 29110-1:2016, 3.74, düzəliş - Qeyd1-in yalnız sonuncu cümləsi saxlanılıb.]

3.1.3

ümumiləşdirmə

<maşın öyrənməsi> kontekstində – təlim keçmiş modelin əvvəllər rast gəlmədiyi giriş verilənləri əsasında dəqiq proqnozlar vermək bacarığı

Qeyd 1: Yaxşı ümumiləşdirmə aparan maşın öyrənməsi modeli əvvəllər rast gəlmədiyi giriş verilənlərindən istifadə edərək məqbul dəqiqliklə proqnoz verən modeldir.

Qeyd 2: Ümumiləşdirmə həddindən artıq öyrənmə ilə sıx bağlıdır. Həddindən artıq öyrənmə keçmiş maşın öyrənməsi modeli təlim verilənlərinə çox dəqiq uyğunlaşdırıldığı üçün yaxşı ümumiləşdirilmə apara bilməyəcək.

3.1.4

həddən artıq öyrənmə izafi öyrənmə

<maşın öyrənməsi> kontekstində – təlim verilənlərinə çox dəqiq uyğun olan, lakin yeni verilənlər əsasında ümumiləşdirmə apara bilməyən modelin qurulması

Qeyd 1: Həddən artıq öyrənmə təlim keçmiş model təlim verilənlərindəki mühüm olmayan xüsusiyyətlərdən (yəni ümumiləşdirmələri faydalı nəticələrə səbəb olmayan xüsusiyyətlərdən) istifadə etdikdə, təlim verilənləri həddən artıq küylü (məsələn, çoxsaylı sapmalar) olduqda və ya model təlim verilənləri üçün çox mürəkkəb olduğu hallarda baş verə bilər.

Qeyd 2: Həddən artıq öyrənmə təlim verilənləri üzrə və ayrıca test və yoxlama verilənləri üzrə ölçülən xətalar arasında kifayət qədər fərq olduqda aşkar edilə bilər. Təlim verilənləri ilə produksiya verilənləri ("production data") arasındakı kifayət qədər böyük uyğunsuzluq həddən artıq öyrənmə keçmiş modellərin məhsuldarlığına xüsusi təsir göstərir.

3.1.5

yetərsiz öyrənmə

<maşın öyrənməsi> kontekstində – təlim verilənlərinə kifayət qədər uyğun olmayan və yeni verilənlər üzrə qeyri-dəqiq (yanlış) proqnozlar verən modelin qurulması

Qeyd 1: Xüsusiyyətlər düzgün seçilmədikdə, təlim müddəti kifayət qədər olmadıqda və ya böyük həcmli təlim verilənləri əsasında öyrədilməsi baxımından məhdud imkanlara (yəni ifadə gücünə) malik model çox sadə olduqda öyrənmə prosesi yetərsiz ola bilər.

3.2 Alətlər

3.2.1

geri yayılma

neyron şəbəkəsinin çıxış layındakı xətdən istifadə edərək ardıcıl əvvəlki laylardakı əlaqələrin çəki əmsallarını tənzimləyən və optimallaşdıran təlim metodu

3.2.2

öyrənmə dərəcəsi

gradiyent metodu üçün addım ölçüsü

Qeyd 1: Öyrənmə dərəcəsi modelin optimal həllə yaxınlaşmasını və bu yaxınlaşmanın tezliyini müəyyən edir ki, bu da onu neyron şəbəkələrinin sazlanması üçün mühüm hiperparametrdir.

3.3 Verilənlər

3.3.1

sinif

Verilənlər çoxluğunun bir hissəsi olan və ümumi atributlara malik elementlərin insan tərəfindən müəyyən edilən kateqoriyası

MİSAL: "telefon", "masa", "stul" və "tennis topu" siniflərə nümunələrdir. İş masası, yemək masası, dərs masası, qəhvə masası "masa" sinfinə aiddir.

Qeyd 1: Siniflər adətən məqsəd dəyişənləridir və onlara ad verilir.

3.3.2

klaster

Verilənlər çoxluğunun bir hissəsi olan və ümumi atributlara malik elementlərin avtomatik yaradılan kateqoriyası

Qeyd 1: Klasterlərin adının olmasına ehtiyac yoxdur.

3.3.3

xüsusiyyət

<maşın öyrənməsi> kontekstində – obyekt və ya hadisənin xassələr çoxluğuna nəzərən ölçülə bilən xassəsi

Qeyd 1: Xüsusiyyətlər təlimlərdə və proqnozlaşdırmada xüsusi rol oynayır.

Qeyd 2: Xüsusiyyətlər müvafiq obyektlərin müəyyən edilməsi üçün "maşın tərəfindən oxuna bilən" üsul təqdim edir. Alqoritmin obyektlərlə və ya hadisələrlə bilavasitə işləməyəcəyini nəzərə alaraq, xüsusiyyətlərin təsviri bütün faydalı məlumatları ehtiva etməlidir.

3.3.4
distansiya
məsafə

<maşın öyrənməsi> kontekstində – fəzada iki nöqtə arasında ölçülən yaxınlıq

Qeyd 1: Maşın öyrənməsində adətən Evklid və ya “düz xətt” məsafəsi istifadə olunur.

3.3.5
nişanlanmamış

məqsəd dəyişənini ehtiva etməyən nümunənin xassəsi

4 QISALTMALAR

AI	artificial intelligence	süni intellekt
API	application programming interface	tətbiqi proqramlaşdırma interfeysi, tətbiqi proqram interfeysi
AUC	area under the curve	əyri altındakı sahə
BM	Boltzmann machines	Bolsman maşınları
CapsNet	capsule neural network	kapsul neyron şəbəkəsi
CG	conjugate gradient	qoşma qradiyent
CNN	convolutional neural network	konvolyusiyalı (bürünmə) neyron şəbəkəsi
DBN	deep belief networks	dərin etimad şəbəkəsi
DCNN	deep convolutional neural network	dərin konvolyusiyalı (bürünmə) neyron şəbəkəsi
FFNN	feed forward neural network	irəli yayılma neyron şəbəkəsi
FNR	false negative rate	yanlış mənfi hallar tezliyi
FPR	false positive rate	yanlış müsbət hallar tezliyi
GRU	gated recurrent unit	qapılı (ventilli) rekkurent şəbəkə
LSTM	long short-term memory	uzun qısamüddətli yaddaş
MAE	mean absolute error	orta mütləq səhv
MDP	Markov decision process	Markov qərar qəbuletmə prosesi
ML	machine learning	maşın öyrənməsi

NN	neural network	neyron şəbəkəsi
NNEF	neural network exchange format	neyron şəbəkəsi mübadiləsi formatı
NPV	negative predictive value	mənfi proqnoz ehtimalı
ONNX	open neural network exchange	açıq neyron şəbəkə mübadiləsi (standartı)
PCA	principal component analysis	əsas komponent təhlili
PHI	personal or protected health information	fərdi məlumatlar və ya səhhətə dair mühafizə olunan məlumatlar
PII	personally identifiable information	şəxsi identifikasiya məlumatları, fərdi məlumatlar
PPV	positive predictive value	müsbət proqnoz ehtimalı
REST	representational state transfer	reprezentativ vəziyyətlər transferi
RNN	recurrent neural network	rekurrent neyron şəbəkəsi
ROC	receiver operating characteristics	qəbul edicinin iş xüsusiyyətləri
SGD	stochastic gradient descent	stoxastik qradient enişi
SVM	support vector machine	dəstək vektoru maşını
TNR	true negative rate	doğru mənfi hallar tezliyi
TPR	true positive rate	doğru müsbət hallar tezliyi

5 İCMAL

ISO/IEC 22989 beynəlxalq standartına əsasən, maşın öyrənməsi (ML) model parametrlərinin hesablama üsulları ilə optimallaşdırılması prosesi kimi müəyyən olunur, belə ki, modelin davranışı verilənləri və ya təcrübəni əks etdirir. Neyronların (neyron şəbəkələrinin) modelləşdirilməsi və verilənlər əsasında öyrənə bilən kompüter proqramlarının işlənməsi hələ 1940-cı illərin əvvəllərindən başlayaraq araşdırma mövzusu olmuşdur. Maşın öyrənməsi (ML) durmadan genişlənən bir sahədir və bir çox sənaye sektorunda yeni tətbiqlərin meydana gəlməsi ilə müşayiət olunur. Bu irəliləyişin reallaşması isə böyük həcmli verilənlər və hesablama resurslarının mövcudluğu sayəsində mümkün olur. ML metodlarına neyron şəbəkələri və dərin öyrənmə daxildir.

ISO/IEC 22989 standartında süni intellekt ekosistemi funksional səviyyələr baxımından təqdim olunur və maşın öyrənməsi (ML) bu Sİ ekosisteminin mühüm komponentidir. Şəkil 1-də modelin komponentləri, proqram təminatı vasitələri və texnikaları, eləcə də verilənlərdən ibarət olan ML sistemi əks olunur.

Bu standartın 6-cı bölməsində ML sisteminin ayrı-ayrı komponentləri daha ətraflı şəkildə təsvir edilir.

Standartın 7-ci bölməsində müxtəlif ML yanaşmaları, eləcə də bu yanaşmaların təlim verilənlərindən asılılığı göstərilir.

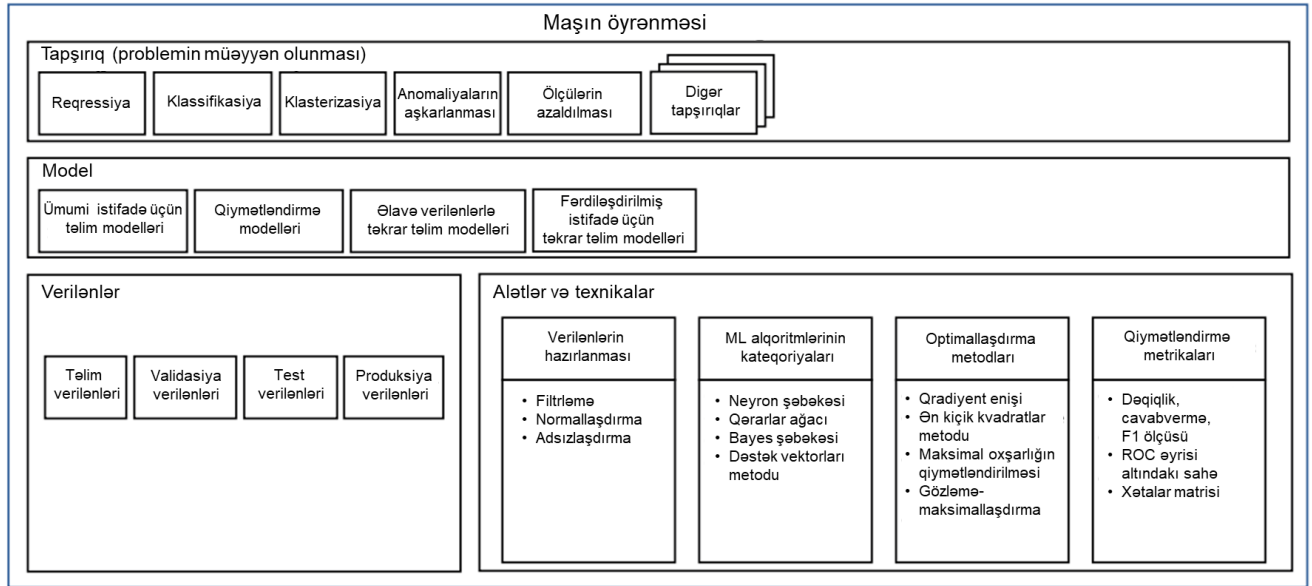
Standartın 8-ci bölməsində ML konveyeri ("pipeline") – ML modelinin işlənməsi, quraşdırılması və istismarı ilə bağlı proseslər izah edilir.

6.5-ci bənd və 7-ci bölmə sənədin qalan hissəsindən fərqli olaraq daha texniki xarakter daşıdığından, bu kontentin oxucu tərəfindən daha yaxşı anlaşılması üçün dərin texniki biliklərə ehtiyac ola bilər.

6 MAŞIN ÖYRƏNMƏSİ SİSTEMİ

6.1 Ümumi müddəalar

Şəkil 1-də ML sisteminin elementləri əks olunur. Burada elementlərin rolları və onların maşın öyrənməsinə aid konkret funksiyaları ümumi şəkildə təsvir edilir. Bu rol və funksiyalar müxtəlif qurumlar (məsələn, ayrı-ayrı təchizatçılar) tərəfindən icra oluna bilər. Şəkil 1 nümunələrin tam siyahısını əks etdirmir. Şəkil 1-in hər bloku üzrə izah xarakterli əlavə məlumatlar 6-cı bölmədə təqdim olunur.



Şəkil 1 - ML sisteminin elementləri

Şəkil 1-də modelin qurulması və istifadəsi üzrə aləmlərləri çoxsəviyyəli yanaşma kimi nəzərdən keçirmək olar (yeni tətbiqlər tapşırıqları həll etmək üçün istifadə olunan modellər əsasında hazırlanır). Modelin qurulması və istifadəsi isə proqram təminatı vasitələri və texnikaları, eləcə də verilənlərdən asılıdır.

Bir ML sistemi bir neçə ML modelinin kombinasiyasından ibarət ola bilər. Sistem komponentlərini onların giriş və çıxış verilənləri, eləcə də təyinat və ya funksiyası baxımından təsvir etmək mümkündür. Həmçinin bu komponentləri müstəqil şəkildə test etmək də olar.

ML modelləri (quraşdırıldıqdan sonra) proqnoz və ya qərarlar kimi çıxış verilənləri generasiya edir. İlk təlim keçmiş model - əldə olunan zaman artıq təlim keçmiş ML modelidir. Bəzi hallarda qurulmuş model müxtəlif sahələrdə analoji tapşırıqlar üçün tətbiq etmək olar. Transfer öyrənmə digər müvafiq tapşırığı yerinə yetirmək üçün ilk təlim keçmiş ML modelini modifikasiya etmək üçün texnikadır.

Bu standartda tətbiq dedikdə, həm bir və ya bir neçə ML modelinin təyinatlı istifadəsi, həm də həmin istifadəni həyata keçirən konkret proqram təminatı nəzərdə tutulur. Bir qayda olaraq, tətbiqlər yaratmaq üçün ML modelləri digər proqram təminatı komponentlərinə inteqrasiya olunur. ML-dən istifadə edən tətbiqlər emal etdikləri giriş verilənlərinin və yerinə yetirdikləri tapşırıqların növlərinə görə fərqlənir. Bəzi tətbiqlərdə ML yüksək səviyyəli proqnoz və ya qərarlar generasiya edir, digər tətbiqlərdə isə dar ixtisas sahəsini əhatə edən məsələləri həll edir.

Giriş verilənləri və tapşırıqlardakı fərqlər, eləcə də quraşdırılma variantları, dəqiqlik və etibarlılıq kimi amillər tətbiqlərin fərqli layihələndirilməsi (arxitektura) ilə nəticələnir. Sİ tətbiqləri xüsusi fərdiləşdirilmiş layihələndirmələrdən və ya tətbiq sahəsinə aid standart layihələndirmə nümunələrindən istifadə edə bilər.

Tətbiqin məntiqi giriş verilənlərinin formatı, çıxış verilənləri və bəzi hallarda istifadə olunan ML modelləri arasında verilənlərin transformasiyası və axını ilə müəyyən olunur. Bütün hallarda, ML alqoritmlərinin və verilənlərin hazırlanması texnikalarının seçimi tətbiqin tapşırıqlarına uyğunlaşdırılır.

6.2 Tapşırıq

6.2.1 Ümumi müddəalar

“Tapşırıq” dedikdə, konkret hədəfə çatmaq üçün tələb olunan tədbirlər nəzərdə tutulur. ML-də isə bu termin ML modelindən istifadə etməklə həll olunacaq məsələnin müəyyənləşdirilməsini ehtiva edir. ML tətbiqi üçün bir və ya bir neçə ML tapşırığı müəyyən edilə bilər. Tapşırıqların həlli üçün addımlar çoxluğu kimi təqdim olunan və proqram kodu vasitəsilə həyata keçirilən hər hansı funksiyadan istifadə etmək əvəzinə, təlim keçmiş ML modelini produksiya verilənlərinə tətbiq etmək olar. Faktiki olaraq, təlim keçmiş ML modeli tapşırığı həll etmək üçün proqramçı tərəfindən yazıla biləcək hipotetik funksiyanın aproksimasiyası olan məqsəd funksiyasını həyata keçirir.

ML tapşırığının qoyuluşu məsələnin, eləcə də verilənlərin formatının və xüsusiyyətlərinin müəyyən olunmasını ehtiva edir.

Aşağıdakı altbəndlərdə təsvir olunan tapşırıqların siyahısı tam deyil və sadəcə nümunə üçün verilir.

6.2.2 Reqressiya

Reqressiya tapşırıqları təlim verilənləri çoxluğuna ən uyğun modeli öyrətməklə kəsilməz funksiyanın qiymətlərinin proqnozlaşdırılmasını ehtiva edir. Reqressiya tapşırığında təlim keçmiş reqressiya modeli fərdiləşdirilmiş istifadəçi fəzasını ("custom space") təmsil edir. Təlim keçmiş model yeni produksiya verilənləri nümunələrinə tətbiq edildikdə, bu nümunələr təlim keçmiş reqressiya modeli ilə müəyyən olunan fərdiləşdirilmiş istifadəçi fəzasına inikas olunur.

Reqressiya, əsasən, real prosesin ədədlə ifadə olunan qiymətlərini eyni proseslə bağlı əvvəlki ölçmə və ya müşahidələr əsasında proqnozlaşdırmaq üçün istifadə olunur. Reqressiyadan istifadə halları aşağıdakıları ehtiva edir:

- səhm bazarında qiymətlərin proqnozlaşdırılması;
- yayımlanan videoların izləyici auditoriyasının yaşının proqnozlaşdırılması;
- müxtəlif klinik ölçülərə əsasən, orqanizmdə prostat spesifik antigenin miqdarının proqnozlaşdırılması.

6.2.3 Klassifikasiya

Klassifikasiya tapşırıqları giriş verilənləri nümunələrinin müəyyən kateqoriya və ya sinfə aid edilməsini proqnozlaşdırmaqdan ibarətdir. Klassifikasiya binar (yəni doğru və ya yanlış), çoxsınıflı (yəni mümkün imkanlardan biri) və ya çoxqiymətli (yəni mümkün imkanların istənilən sayı) ola bilər. Məsələn, təsvirdəki obyektin nəinki pişik və ya it, hətta tam fərqli bioloji növdən olduğunu proqnozlaşdırmaq üçün klassifikasiyadan istifadə oluna bilər.

Siniflər adətən diskret və nizamlanmamış çoxluqlardan seçilir, ona görə də bu problemə reqressiya tapşırığı kimi baxıla bilməz. Məsələn, bir neçə simptomun tibbi diaqnozu insult, aşırı dərman dozası, qıcolma ola bilər ki, burada siniflər nizamlanmayıb və bir sinifdən digərinə kəsilməz keçid yoxdur.

Aşağıdakılar klassifikasiyanın istifadəsinə aiddir:

- Sənədlərin bir neçə sinif üzrə qruplaşdırıldığı sənəd klassifikasiyası və e-poçt spamının filtrlənməsi. Məsələn, spam filtrində nümunələrin klassifikasiyası üçün “spam” və “spam deyil” kimi iki sinifdən istifadə olunur;
- Nümunələrin növlər üzrə klassifikasiyası. Məsələn, ML klassifikasiya modeli kasa yarpağının uzunluq və eni, eləcə də ləçəyin uzunluq və eni barədə məlumatlar əsasında çiçəyin növünü proqnozlaşdırmağa bilər;
- Təsvirlərin klassifikasiyası. Təsvirlər çoxluğunu (məsələn, mebel şəkillərini) götürsək, həmin təsvirlərdə göstərilən obyektləri tanımaq və adlandırmaq üçün ML sistemindən istifadə oluna bilər.

6.2.4 Klasterizasiya

Klasterizasiya tapşırıqları giriş verilənləri nümunələrinin qruplaşdırılmasını ehtiva edir. Klassifikasiya tapşırıqlarından fərqli olaraq, klasterizasiya tapşırıqlarında siniflər əvvəlcədən yox, klasterizasiya prosesində kimi müəyyən edilir. Klasterizasiya oxşar (homogen) verilənləri müəyyən etmək məqsədilə verilənlərin hazırlanması addımı kimi istifadə edilə bilər. Bu homogen verilənlərdən daha sonra müəllimlə (supervayzerlə) məşin öyrənməsində təlim verilənləri kimi istifadə etmək olar. Klasterizasiya digər nümunələrə bənzəməyən giriş verilənləri nümunələrini müəyyən edərək sapmaların və ya anomaliyaların aşkarlanması üçün də istifadə edilə bilər. Klasterizasiya tapşırıqlarına dair tətbiq nümunələrinə faylların çeşidlənməsi və ya nizamlanması aiddir.

6.2.5 Anomaliyaların aşkarlanması

Anomaliyaların aşkarlanması gözlənilən qanunauyğunluqlara uyğun gəlməyən giriş verilənlərinin nümunələrini müəyyən etməkdən ibarətdir. Anomaliyaların aşkarlanması dələduzluğu və ya fərqli aktivliyi aşkar edən tətbiqlər üçün faydalı ola bilər.

Anomaliyaların aşkarlanması üçün ML modeli giriş verilənləri nümunələrinin verilmiş paylanma üçün tipik olub-olmadığını proqnozlaşdırır.

6.2.6 Ölçülərin azaldılması

Ölçülərin azaldılması faydalı məlumatların əksəriyyətini saxlamaqla konkret nümunələr üzrə atributların (ölçülərin) sayını azaltmaqdan ibarətdir.

Ölçülərin azaldılması verilənlər çoxluğunun ən faydalı xüsusiyyətlərinə önəm verməyə və bununla da hesablama xərclərini aşağı salmağa imkan yarada bilər.

Ölçülərin azaldılması “ölçü lənəti” kimi tanınan çox sayda xüsusiyyətlərin saxlanması müxtəlif mənfi təsirlərini aradan qaldırır. Ölçülərin azaldılması verilənlərin araşdırılması və modelin analizi üçün də faydalıdır.

Ölçülərin azaldılması metodlarına müəllimlə, müəllimsiz və ya qismən müəllimlə öyrənmə daxildir [1].

6.2.7 Digər tapşırıqlar

Müxtəlif məqsədləri və gözlənilən nəticələri olan çox sayda digər tapşırıqlar da mövcuddur. Müəyyən tətbiq üçün tipik olan tapşırıqlar ola bilər. Digər tapşırıqlara misal olaraq mətn və ya təsvirlərin semantik seqmentləşdirilməsi, məşin tərcüməsi, nitqin tanınması və ya sintezi, obyektin lokallaşdırılması və təsvirlərin generasiyasını göstərmək olar.

Planlaşdırma baxımından, tapşırıq – ətraf mühitin vəziyyətini müşahidə etməklə, agent(lər)in hərəkət ardıcılığının (yerinə yetirilməli olduğu tədbirlərin) optimallaşdırılması deməkdir.

Tapşırıqların müxtəlifliyinə baxmayaraq, digər tapşırıqlardan bəziləri arasında əlaqə yaratmaq üçün bir sıra konsepsiyalar formalaşdırılıb. Belə konsepsiyalardan biri də modelin gözlənilən nəticəsinin vahid qiymətdən fərqli olaraq strukturlaşdırılmış obyekt olduğu tapşırıqlara uyğun strukturlaşdırılmış proqnozlaşdırmaadır.

Strukturlaşdırılmış proqnozlaşdırma alınmış nəticədə qanunauyğunluqları (ya aşkar şəkildə modelləşdirməklə, ya da qanunauyğunluqları daxilən modelləşdirən model sayəsində bütün strukturu proqnozlaşdırmaqla) nəzərə alan hesablama metodlarını tələb edir.

Strukturlaşdırılmış proqnozlaşdırma üzrə istifadə halları aşağıdakıları ehtiva edir:

- təbii dildəki cümlələr üçün sintaktik təhlil ağacının qurulması;
- bir dildəki cümlənin başqa dildəki cümləyə tərcümə edilməsi;
- zülal strukturunun proqnozlaşdırılması;
- təsvirin semantik seqmentasiyası.

6.3 Model

ISO/IEC 22989:2022 (AZE ISO/IEC 22989:2023) beynəlxalq standartının 3.3.7-ci bəndində ML modelinə giriş verilənlərinə əsaslanaraq nəticə və ya proqnozlar generasiya edən riyazi konstruksiya kimi tərif verilir. ML modeli (hər ikisi seçilmiş ML alqoritmi əsasında müəyyən olunan) verilənlərin strukturu və bu strukturu emal etmək üçün proqram təminatından ibarətdir. Model verilmiş məsələnin həlli üçün vacib olan giriş və çıxış verilənləri ilə konfigurasiya edilir.

Model təlim verilənlərinin müvafiq statistik xüsusiyyətlərini təmsil etmək üçün təlim keçir. Faktiki olaraq, model əldə edilmiş bu biliyin sonradan real praktiki sahədə tətbiqi üçün təlim prosesində təlim verilənləri vasitəsilə problemin həlli yolunu “öyrənir”.

ML modelləri optimal həllərin approksimasiyası olan nəticələr hasil edir. ML alqoritmləri bu approksimasiyanı həyata keçirmək üçün statistik optimallaşdırma metodlarından istifadə edir. Bu metodların tətbiqi nəticəsində modelin giriş verilənlərinin çıxış verilənlərinə inikası təlim verilənləri əsasında öyrənilən qanunauyğunluqları (obrazları) əks etdirir. Obrazlar korrelyasiya, səbəb-nəticə əlaqələri və ya verilənlər obyektlərinin kateqoriyalarına aid ola bilər. ML modelləri istifadə olunan təlim verilənlərinin nəticəsində formalaşır. Beləliklə,

istifadə olunan verilənlər natamamdırsa və ya özünəməxsus sosioloji qərəzi əks etdirirsə, modelin məhsuldarlığı da bunu əks etdirəcək. Buna görə də təlim modelləri üçün istifadə olunan verilənlər çoxluğuna diqqət yetirilməlidir. Maşın öyrənməsi prosesində yaradılan və təlim keçmiş model tərəfindən təqdim olunan məntiq proqramçı tərəfindən müəyyən edilmir, təlim prosesi zamanı formalaşır.

Modelin məhsuldarlığını müəyyənləşdirmək üçün qiymətləndirmə metrikalarından istifadə edilir.

Təkrar təlim prosesi təlim keçmiş modeli müxtəlif təlim verilənləri ilə öyrətməklə yeniləməkdən ibarətdir. Təkrar təlimə kifayət qədər çox təlim verilənlərinin olmaması, verilənlərin dreyfi ("drift") və konsepsiya dreyfi də daxil olmaqla, bir sıra səbəblər baxımından ehtiyac yarana bilər.

Verilənlərin dreyfi zamanı modelin proqnozlarının dəqiqliyi produksiya verilənlərinin statistik xüsusiyyətlərindəki dəyişikliklər səbəbindən zamanla azalır. Bu halda modelin produksiya verilənlərini daha yaxşı təmsil edən yeni təlim verilənləri ilə təkrar öyrədilməsi vacibdir.

Konsepsiya dreyfində qərar qəbuletmənin sərhədləri dəyişir, bu da öz növbəsində hətta verilənlərin dəyişməsinə baxmayaraq, proqnozların dəqiqliyini aşağı salır. Konsepsiya dreyfi zamanı təlim verilənlərində məqsəd dəyişənləri yenidən nişanlanmalı və model təkrar təlim keçməlidir.

Təkrar təlim, həmçinin transfer öyrənmə və ML modelinin optimallaşdırılması və ya dəyişdirilməsi kimi məqsədlər üçün də həyata keçirilə bilər.

Bəzi modellər əlçatanlığı ilə fərqlənir, onlar müəyyən istifadə halları üçün təkrar təlimləmə və ya optimallaşdırıla, yaxud bu modellərdən olduğu kimi istifadə edilə bilər. Buna misal olaraq kommersiya təyinatlı maşın tərcüməsi modelini qeyd etmək olar. Hüquqi sənədlərin tərcümə edilməsi üçün bu modelə təkrar təlim keçmək mümkündür.

Fasiləsiz öyrənmə təkrar təlimin xüsusi halıdır, belə ki, produksiya verilənləri ilə daimi təlim nəticəsində modelin məhsuldarlığı davamlı olaraq inkişaf edir. Belə hallarda modelin fəaliyyətinin fasiləsiz monitorinqinə və ya modelin məqbul davranışını müəyyən edən "məhdudlaşdırıcılar"ın ("guard-rails") tətbiq edilməsinə ehtiyac yarana bilər.

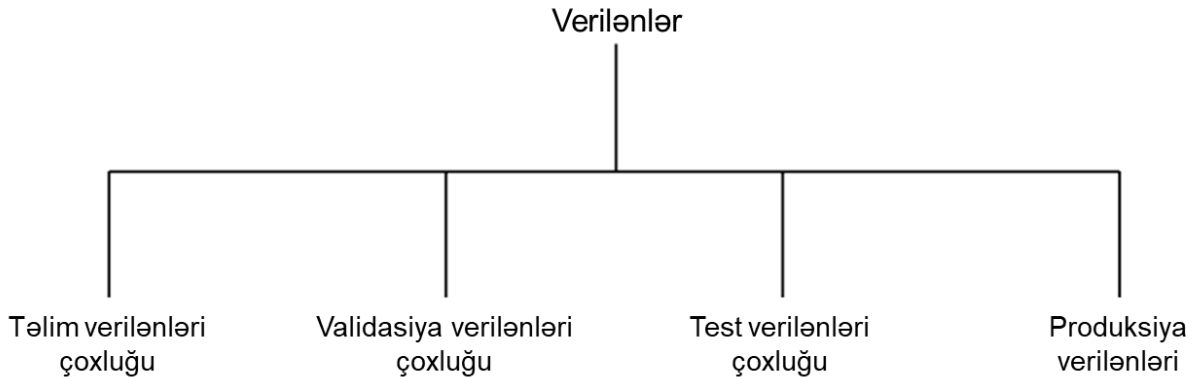
6.4 Verilənlər

Şəkil 2-də verilənlər konsepsiyasının bir-birini istisna edən dörd kateqoriyaya bölündüyünü əks etdirən diaqram təqdim olunur:

- a) Namizəd-modellərin parametrlərini təyin etmək üçün istifadə olunan təlim verilənləri çoxluğu;
- b) Sİ sahəsindən asılı olaraq (məsələn, təbii dilin emalında), performans meyarına uyğun olaraq ən yaxşı modelin seçilməsi üçün istifadə edilən həqiqiliyin yoxlanılması (bundan sonra – validasiya) verilənləri çoxluğu. Validasiya verilənləri çoxluğu işlənilmə verilənləri çoxluğu ("development dataset") da adlandırılır;
- c) Modelin ümumiləşdirmə qabiliyyətini yoxlamaq və gələcək verilənlər üzrə effektivliyini müəyyən etmək üçün istifadə olunan test verilənləri çoxluğu;

d) Model tərəfindən proqnozlaşdırma üçün istifadə ediləcək istismar verilənlərindən ibarət produksiya verilənləri. Produksiya verilənlərinin paylanması təlim, validasiya və test verilənləri çoxluğunun paylanmasından fərqlənə bilər.

Təlim, validasiya və test verilənləri çoxluğu imitasiya və perturbasiya verilənləri (küylü verilənlər) kimi süni verilənlərlə tamamlana bilər.



Şəkil 2 – Anlayışlar diaqramı: verilənlər və verilənlər çoxluqları

Bu müxtəlif növ verilənlər ya yalnız giriş verilənlərindən, ya da nişanlarla əlaqələndirilmiş giriş verilənlərindən (gözlənilən çıxış verilənləri) ibarət ola bilər. Validasiya və test verilənləri çox vaxt nişanlansa da, produksiya verilənləri bir qayda olaraq nişanlanmır. Təlim verilənləri xüsusunda isə bu məsələ ML yanaşmasından asılıdır, belə ki, təlim verilənləri həm nişanlanmamış, həm qismən və ya tam nişanlanmış ola bilər.

Nişanlanmış təlim verilənləri ML alqoritminə giriş dəyişənləri ilə məqsəd dəyişəni arasındakı statistik əlaqələri müəyyən etməyə imkan verir. Nişanlanmamış təlim verilənləri sayəsində isə ML alqritmi giriş verilənlərinin statistik korrelyasiyasını və strukturunu müəyyən edə bilər.

6.5.5 yarımbəndində müzakirə olunduğu kimi, validasiya və test verilənlərinin hər ikisi statistik fəaliyyət göstəriciləri üçün istifadə olunur, lakin onların istifadəsi fərqlidir. Belə ki, validasiya verilənləri hiperparametrlərin tənzimlənməsini, test verilənləri isə modelin qiymətləndirilməsini nəzərdə tutur. Test verilənlərinin məqsədi təlim keçmiş modelin produksiya verilənləri əsasında yaxşı nəticə göstərəcəyini və ya ümumiləşdirmənin təmin olunmasını yoxlamaqdır. Lazımı şəkildə ümumiləşdirmə apara bilməyən modellər (təlim verilənlərinə münasibətdə) “həddən artıq təlim keçmiş” modellər adlanır.

Nəzərə almaq lazımdır ki, “test verilənləri” termininin istifadəsi ML proseslərinin spesifikliyi ilə məhdudlaşır və bu termin ML komponentlərindən istifadə edən inteqrasiya olunmuş sistemin verifikasiyası və validasiyası kontekstindəki istifadəsindən fərqlənir, belə ki, bu halda sözügedən termin (ML ilə konkret əlaqəsi olmadan) verifikasiya və validasiya üçün istifadə edilən istənilən verilənləri əhatə edə bilər. Həmçinin nəzərə almaq lazımdır ki, “validasiya verilənləri” termininin ML ilə bağlı spesifik istifadəsi (modelin qurulmasına aid olduğu üçün) verifikasiya və validasiya ilə əlaqəli deyil.

ML proseslərinin etibarlı tətbiqi üçün təlim, validasiya və test verilənləri fərqli olmalıdır (bu verilənlər çoxluqları kəsişməməlidir). Təlim, validasiya və test çoxluqları eyni verilənlər çoxluğunun bölünməsi yolu ilə, yaxud ayrıca əldə edilə bilər. İdeal konfigurasiya şəraitində

onların hamısı eyni statistik paylanmaya malik olur, lakin istifadə variantından və ML yanaşmasından asılı olaraq, fərqli paylanmaya zərurət yarana bilər.

Təlim keçmiş modelin etibarlı qiymətləndirilməsi üçün test verilənləri produksiya verilənlərinin paylanmasına mümkün qədər yaxın paylanmaya malik olmalıdır.

Model quraşdırıldıqdan sonra produksiya verilənləri onun üçün əlçatan olur. modelin dəqiq proqnozlar verməsi üçün, ümumilikdə, produksiya verilənlərinin paylanması təlim və validasiya verilənlərinin paylanmasına analoji olmalıdır. Buna baxmayaraq, paylanmalarda fərqlilik olduqda, problemləri aradan qaldıra biləcək xüsusi üsullar mövcuddur.

Zaman keçdikcə, produksiya verilənlərinin paylanması dəyişə ("drift") bilər ki, bu da modelin yeni verilənlər əsasında təkrar təlim keçməsinə tələb edir. Modelin produksiya verilənlərindəki yeni qanunauyğunluqlara dinamik şəkildə adaptasiyası zəruri olduğu hallarda, produksiya verilənlərindən əldə olunan məlumatlardan faydalanmaqla, model fasiləsiz olaraq təkrar təlim keçə bilər. Belə hallar 6.3-cü bənddə müzakirə olunur.

6.5 Alətlər

6.5.1 Ümumi müddəalar

ML modelinin qurulması üçün verilənlərin hazırlanması, ML alqoritmləri, optimallaşdırma metodları və qiymətləndirmə metrikaları kimi kateqoriyalara bölünən alətlərdən istifadə edilir. ML modelinin effektivliyi qiymətləndirmə metrikaları generasiya edən alətlər vasitəsilə qiymətləndirilir.

ML modelinin qurulması hesablama sistemlərinə olan tələblər və böyük təlim verilənləri çoxluğunun istifadəsi səbəbindən çox vaxt yüksək məhsuldarlıqlı hesablama resursları (iş yükü) tələb edir. ML modellərinin işlənməsi və təlim keçməsinə hesablama sistemlərinin və məlumatların saxlanması sistemlərinin məhsuldarlığı da təsir göstərə bilər.

ML ilə bağlı fundamental problemlərə statistik təhlil, alqoritmlərin layihələndirilməsi və optimallaşdırma daxildir. Statistik təhlil təlim verilənlərindən əsasən əldə edilən riyazi modellərin prinsiplərini ehtiva edir. Alqoritmlərin layihələndirilməsi ML modelini qurmaq üçün istifadə olunan alqoritmik texnikaların tətbiqi qaydasıdır. Hər hansı ML modelinin optimallaşdırılması da ML-in tətbiqi baxımından mühüm əhəmiyyət kəsb edir. Digər bir problem isə ML-in potensial və imkanlarını başa düşməkdir. ML verilənlərə əsaslandığı üçün o, bir çox hallarda verilənlərə məxsus mövcud çatışmazlıq və qeyri-obyektivliyi təkrarlayacaq, gücləndirəcək və sürətləndirəcəkdir.

6.5.2 Verilənlərin hazırlanması

Verilənlərin hazırlanması barədə məlumat 8.3-ci bənddə verilir.

6.5.3 Maşın öyrənməsi alqoritmlərinin kateqoriyaları

6.5.3.1 Ümumi müddəalar

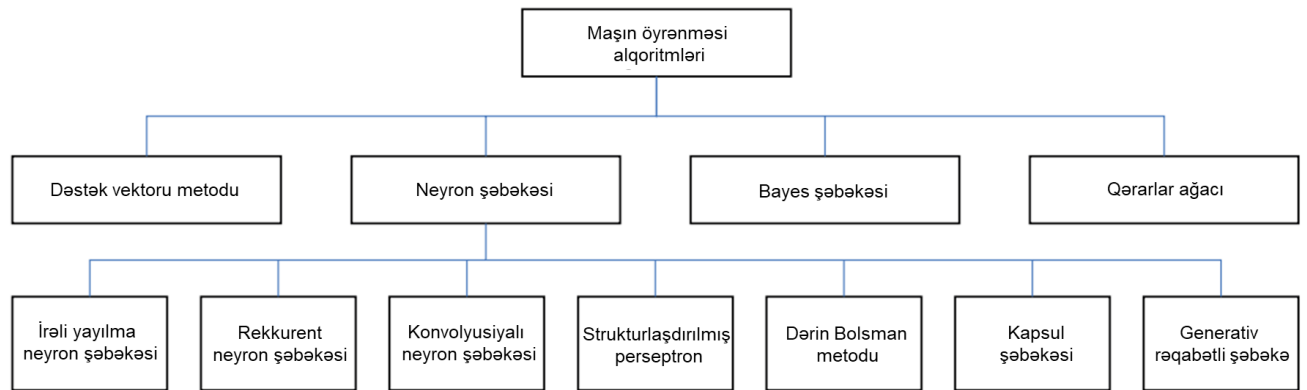
ML alqoritminin seçimi ML modelinin hesablama strukturunu və təlimlənməsi üçün yanaşmanı müəyyən edir.

Alqoritmlər aşağıdakılar daxil olmaqla ML-in müxtəlif məqsədləri üçün istifadə edilə bilər:

- verilənlərin hazırlanma mərhələsinin bir hissəsi olan informasiyanın təqdimatı alqritmi. Bu, xüsusiyyətlər mühəndisliyini (“feature engineering”) nəzərdə tutur;
- ML modelinin qurulması üçün istifadə olunan alqritm.

ML alqoritmləri ilə ML modelləri arasındakı əlaqəni $y = \theta_0 + \theta_1 x$ birdəyişənli xətti funksiyasının həllini nəzərə almaqla təsvir etmək olar. Burada y - çıxış verilənləri və ya nəticə, x - giriş verilənləri, θ_0 - düz xətt parçası ($x = 0$ olduqda y -in qiyməti), θ_1 isə çəki əmsəlidir. ML-də xətti funksiya üçün θ_0 və θ_1 təyin edilməsi prosesi xətti reqressiya kimi tanınır. Əgər ($y = \theta_0 + \theta_1 x$) birdəyişənli xətti funksiya xətti reqressiyadan istifadə edərək təlim keçibsə, nəticədə $y = 3 + 7x$ modeli qurula bilər.

Şəkil 3-də ML alqoritmlərinin müxtəlif kateqoriyalarına dair nümunələr təsvir olunur.



Şəkil 3 - ML alqoritmlərinin müxtəlif kateqoriyalarına dair nümunələr

Əksər hallarda, ML modelinin strukturunu və təlimlənməsi üçün yanaşmanı müəyyən etmək üçün yalnız ML alqritminin seçimi kifayət etmir, belə ki, bir çox alqoritmlər üçün hiperparametrlərin də seçilməsinə ehtiyac yaranır.

Hiperparametrlər öyrənmə prosesinə təsir edən ML alqritminin xarakteristikalarıdır. Hiperparametrlər model parametrlərinin qiymətləndirilməsi üçün həyata keçirilən proseslərdə istifadə oluna bilər. Neyron şəbəkələri üçün hiperparametrlərə misal olaraq şəbəkə laylarının sayını, hər bir layın enini və aktivasiya funksiyasının növünü göstərmək olar. Bütün mümkün kombinasiyalar arasında optimal hiperparametrlər çoxluğunu təyin etmək üçün praktiki yanaşmalardan biri də məhdudiyət funksiyası vasitəsilə təsadüfi axtarışlar aparmaq və validasiya verilənləri çoxluğu əsasında performansını ölçməkdir. Bu addım “model seçimi” və ya “hiperparametrin tənzimlənməsi” adlanır.

6.5.3.2 Neyron şəbəkələri

6.5.3.2.1 Ümumi müddəalar

Neyron şəbəkələri (NN) mürəkkəb problemlərin həlli məqsədilə canlıların müşahidə, öyrənmə, təhlil və qərar qəbulu ilə bağlı intellektual qabiliyyətini imitasiya etməyə çalışır.

Buna görə də, neyron şəbəkələrinin layihələndirilməsi zamanı neyronların həm insan, həm də heyvan beynində əlaqələnməsi prinsipi əsas götürülür. NN sadə hesablama vahidləri olan “neyronlar”ın çoxlaylı şəbəkələrindən istifadə edir. Bu NN-də hər bir hesablama vahidi çıxış qiymətini hesablamaq üçün giriş qiymətləri çoxluğunu ehtiva edir ki, həmin çıxış qiyməti də öz növbəsində “aşağı yerləşən” digər neyronlara ötürülür. NN-in strukturu qarşılıqlı əlaqədə olan emal elementlərindən ibarətdir. Qovşaqlar (və ya neyronlar) şəbəkəsində neyronlar çəki əmsallı tillərlə birləşdirilir. Şəbəkənin ən sadə forması giriş verilənlərinin bir və ya bir neçə xüsusiyyətinə malik bir və ya bir neçə giriş neyronundan, həmçinin bir və ya bir neçə neyronun çıxış layından ibarətdir. Hər bir neyron əvvəlki laydan bir və ya bir neçə girişi qəbul edə bilər, onun çıxışı isə əvvəlki laydakı girişlərin çəki əmsallı kombinasiyasından adətən qeyri-xətti asılı olan aktivləşdirmə funksiyası ilə verilir. NN faktiki nəticə ilə gözlənilən nəticə arasında müqayisə apararaq və çəki əmsallarını tənzimləmək üçün xətdən istifadə etməklə, məlum giriş verilənləri ilə təlim keçərək “öyrənir”.

NN-in çıxış və giriş layları arasında gizli laylar mövcud olarsa, şəbəkə “çoxlaylı perseptron” və ya “çoxlaylı NN” adlanır. “Dərin neyron şəbəkəsi” və ya “dərin öyrənmə” çoxsaylı gizli layı olan neyron şəbəkələrinə aiddir. Dərin öyrənmə çoxsaylı gizli layı olan NN-in təlimlənməsi ilə zəngin iyerarxik təqdimatların yaradılmasına bir yanaşmadır. Bu proses neyron şəbəkəsinə yekun nəticəni tədriclə təkmilləşdirməyə imkan verir. Dərin öyrənmə xüsusiyyətlərinin əldə edilməsinə (xüsusiyyətlər mühəndisliyi) olan ehtiyacı azalda və ya aradan qaldıra bilər, çünki ən uyğun xüsusiyyətlər avtomatik olaraq müəyyən edilir.

NN üç əsas növ üzrə təsnif edilə bilər: diskriminativ, generativ və hibrid (əvvəlki iki növün kombinasiyası).

NN-i konneksionist yanaşma kimi qəbul etmək olar. Konneksionizm adətən sadə hesablama vahidləri olan qarşılıqlı əlaqəli vahidlərin şəbəkəsindən istifadə edir [2]. Şəbəkənin davranışı (yuxarıda təsvir edildiyi kimi) təlim sayəsində əldə edilən əlaqələrin çəki əmsallarının dəyişdirilməsi ilə dəqiqləşdirilə bilər. Şəbəkənin qovşaqları informasiyanı paralel emal edir.

Dərin öyrənmə kifayət qədər böyük vaxt və hesablama resursları tələb edə bilər. Maşın öyrənməsi NN-in fəaliyyəti üçün çox vaxt böyük hesablama gücü və yaddaş tələb edir. Modelin mürəkkəbliyini azaltmaq bu şəbəkələri mobil və quraşdırılmış qurğularda istifadə etmək üçün faydalı ola bilər. İcra vaxtının təsiri və enerji sərfiyyatı minimuma endirilə bilər. Bəzi hallarda isə bu azalma bulud xidmətlərinə əsaslanmadan NN-in hətta mobil qurğularda real vaxt rejimində işləməsinə imkan yaradır. Çəki əmsallarının və ya arxitekturanın “sıxılması” (“compressing”) demək olar ki, eyni performans saxlamaqla, modelin mürəkkəbliyini əhəmiyyətli dərəcədə azalda bilər.

6.5.3.2.2 İrəli yayılma neyron şəbəkələri

İrəli yayılma neyron şəbəkələri (FFNN) ən sadə NN arxitekturalarıdır. Bu şəbəkələr məlumatları giriş layından çıxış layına qədər yalnız bir istiqamətdə yayır. Hər bir konkret lay daxilində neyronlararası əlaqə yoxdur. İki qonşu lay adətən “tam əlaqələndirilmiş” hesab olunur, belə ki, bir laydakı hər bir neyronun sonrakı laydakı hər bir neyronla əlaqəsi var. Hər bir əlaqənin öz çəki əmsalı var. Bir qayda olaraq, FFNN nişanlanmış təlim verilənləri ilə geri yayılma metodundan istifadə etməklə təlim keçir, harada ki, çıxış verilənlərinin gözlənilən qiymətini əldə etmək üçün hər bir nümunə etalon (həqiqi)

qiymətlərlə ("ground truth") əlaqələndirilir. FFNN-in faktiki aldığı nəticə ilə etalon qiymətlər arasındakı fərqi "xəta" deyilir [3]. Geri yayılma metodu əvvəlki ardıcıl laylardan əldə edilən əlaqələrin çəki əmsallarını tənzimləmək məqsədilə çıxış layındakı xətdən istifadəni ehtiva edir. Bu metod çox vaxt qradiyent enişi alqoritmi ilə birgə istifadə olunur [4].

6.5.3.2.3 Rekkurent neyron şəbəkəsi

6.5.3.2.3.1 Ümumi müddəalar

Rekkurent neyron şəbəkələri (RNN) [5] nizamlı ardıcılığa malik ardıcıl giriş verilənləri ilə işləyən NN-dir və burada ardıcılıqdakı girişlərin nizamlı olması vacibdir. Bu cür giriş verilənlərinə misal olaraq səs və video axınları kimi dinamik ardıcılıqları, həmçinin mətn və ya ayrıca təsvirlər kimi statik ardıcılıqları göstərmək olar. Prinsip etibarı ilə RNN bir çox sahələrdə istifadə edilə bilər, çünki vaxt qrafiki olmayan verilənlərin əksər formaları (yəni səs və ya videodan fərqli olaraq) ardıcılıq kimi təqdim oluna bilər. Şəkil və ya mətnin bir sətiri eyni anda bir piksel və ya bir simvol ilə verilə bilər, buna görə də vaxtdan asılı olan çəki əmsalları faktiki olaraq vaxta görə deyil, ardıcılıq üzrə daha əvvəl gələn verilənlər üçün istifadə olunur.

Rekkurent neyron şəbəkələrinin (RNN) keçmiş təlim təcrübəsi əsasında müəyyən olunan cari vəziyyət barədə məlumat saxlamaq xüsusiyyəti var. Giriş verilənləri müəyyən zaman ərzində keçidlər ardıcılığı kimi daxil olur və RNN sonrakı keçidlərdə istifadə etmək üçün hər keçiddən daxil olan məlumatları saxlayır, yeni RNN yaddaşa malikdir. Beləliklə, cari keçid üçün RNN-in bir layındakı neyronları təkcə əvvəlki laydakı neyronlarla çəkili əlaqələrə deyil, həm də əvvəlki keçidlərdəki neyronların giriş verilənlərinə malik olur. RNN nitqin tanınması, maşın tərcüməsi, vaxt sıralarının proqnozlaşdırılması və təsvirin tanınması zamanı geniş istifadə olunur. Ümumiyyətlə, RNN (məsələn, "avtomatik doldurma" kimi) məlumatları optimallaşdırmaq və ya tamamlamaq üçün yaxşı seçimdir.

RNN dəyişən uzunluqlu ardıcıl giriş verilənlərinin emalı və ya dəyişən uzunluqlu ardıcıl çıxış verilənlərinin generasiyası üçün əlverişlidir. Ümumi RNN növlərinə LSTM (uzun-qısamüddətli yaddaş) və LSTM-in daha sadə variantı olan qapılı (idarə olunan) rekkurent şəbəkələr (GRU) daxildir.

6.5.3.2.3.2 Uzun-qısamüddətli yaddaş şəbəkələri

Uzun-qısamüddətli yaddaş (LSTM) şəbəkələri həm daha uzun, həm də daha qısa zaman fərqləri ilə verilənləri yadda saxlamağı tələb edən məsələlər üçün nəzərdə tutulmuş RNN formasıdır ki, bu da LSTM-i uzunmüddətli əlaqələri öyrənmək üçün daha uyğun və yararlı edir. Bu şəbəkələr RNN-də geri yayılma metodu ilə əlaqəli itən qradiyent problemini həll etmək üçün nəzərdə tutulub. LSTM-in hər bir hücrəsi (cell) daxili yaddaşa və gizli vəziyyətə malikdir. RNN təlim zamanı geri yayılma metodundan istifadə edir, lakin bu yanaşma yoxa çıxan qradiyent və ya kəskin böyüyən qradiyent problemi ilə üzləşə bilər [6]. Təlim zamanı RNN-in çəki əmsalları təlimin hər iterasiyasında xəta funksiyasının xüsusi törəməsinə əsasən yenilənir. Törəmənin çox kiçik olduğu hallar RNN-in faktiki olaraq təlim keçə (öyrədilə) bilməməsi ilə nəticələnə və ya törəmənin çox böyük olduğu hallar isə RNN-in qeyri-sabit olması ilə nəticələnə bilər.

LSTM uzunmüddətli asılılıqları öyrənmək üçün nəzərdə tutulur və (yaddaşa malik) hücrə, giriş, çıxış və unutma ventilləri (qapıları) olan neyronların kombinasiyasına əsaslanan

arxitekturaya malikdir. Hər bir neyronun onun əvvəlki vəziyyəti haqqında məlumatla malik yaddaş hücrəsi ilə əlaqəsi var. Bu ventillərin (qapıların) funksiyası məlumatların axınına dayandırmaqla və ya ötürülməsinə imkan verməklə məlumatları qorumaqdan ibarətdir. Giriş ventili əvvəlki laydan nə qədər məlumatın hücrədə saxlanıldığını, çıxış ventili isə bu hücrə haqqında məlumatın hansı hissəsinin növbəti laya keçəcəyini müəyyən edir. Unutma ventili şəbəkənin daxili vəziyyətini “sifirlamağa” imkan verir. Məsələn, əgər giriş verilənləri abzasdan və bu abzasda başlayan yeni cümlədən ibarətdirsə, o zaman şəbəkənin əvvəlki cümlədəki bəzi simvolları unutması zəruri ola bilər. LSTM neyronları ventillərin hər biri üçün fərdi çəki əmsalına malikdir, buna görə də, bu şəbəkə NN-in digər növlərindən daha mürəkkəbdir və nəticədə LSTM təlim keçmək və işləmək üçün daha çox resurs tələb edir. LSTM-in, məsələn, Şekspirsayağı yazmaq və ya musiqi bəstələmək kimi mürəkkəb ardıcılıqları öyrənmə bilmək qabiliyyətinin olması göstərilmişdir.

6.5.3.2.4 Konvolyusiyalı neyron şəbəkəsi

Konvolyusiyalı neyron şəbəkələri (CNN) giriş verilənlərindən faydalı məlumatları filtrləmək üçün ən azı bir konvolyusiya layını ehtiva edən neyron şəbəkələrinin bir növüdür. CNN əsasən təsvirlərin emalı [7] və videoların nişanlanması üçün istifadə olunur, bununla yanaşı, audio və ya mətn kimi digər giriş verilənləri növlərinə də (bəzən rekkurent versiyalardan və ya rekkurent CNN-dən istifadə etməklə) tətbiq edilir. Təsvirlərin emalı üçün CNN-in (əlaqələrdən ibarət) nümunələri heyvanların beyin yarımkürələrinin görmə qabarıqna strukturuna bənzəyir. RNN-dən fərqli olaraq, CNN-in müəyyən layındakı hər bir neyron yalnız əvvəlki laydakı neyronlarla əlaqəli olur və əvvəlki vəziyyətinə dair məlumat almır. Verilmiş neyronun giriş neyronları çoxluğu onun “qəbuledici sahəsi” (“receptive field”) adlanır. CNN üçün adətən neyron şəbəkəsinin konvolyusiyalı laylar və pulinq layları da (“pooling layers”) daxil olmaqla, müxtəlif növ laylardan ibarət olması səciyyəvidir. Konvolyusiyalı neyron şəbəkələri giriş verilənlərinin, məsələn, şəkillər (2-D grid) və ya zaman sıraları (1-D grid) kimi torabənzər topologiyaya malik olmasını nəzərdə tutur. Konvolyusiya layları öz giriş verilənlərini vektor və ya matris (məsələn, ikiölçülü təsvir) kimi qəbul edir və ona konvolyusiya tətbiq edir. Bu konvolyusiya sürüşən vergüllü (nöqtəli) skalyar hasil (“sliding dot product”) və ya çarpaz korrelyasiyadır, harada ki, nüvə (kernel) adlanan vektor və ya matris giriş verilənlərinə tətbiq edilir. Konkret konvolyusiya CNN-də giriş verilənlərindən müəyyən xüsusiyyətləri çıxarmaq üçün nəzərdə tutulur və növbəti lay üçün xüsusiyyətlər xəritəsi hazırlayır. Bu yanaşma CNN-in giriş verilənlərinin xüsusiyyətlərinə nəzərən (onların yerləşməsi baxımından) invariant olmasına imkan verir ki, bu da real şəkillərlə iş zamanı son dərəcə arzuolunandır.

Pulinq layları öz giriş verilənlərinin ölçülərini lokal səviyyədə (az sayda giriş verilənlərinin emalı ilə) və ya global səviyyədə (əvvəlki laydan olan bütün giriş verilənlərinin emalı ilə) azaldır. Pulinq əhəmiyyətsiz detalları effektiv şəkildə filtdən (ələkdən) keçirir. Daha yaxşı nəticə əldə etmək üçün bir neçə konvolyusiya və pulinq layı ardıcıl şəkildə istifadə edilə bilər.

Tam əlaqəli laylardan adətən CNN-in çıxışının yaxınlığında istifadə olunur. CNN-in real tətbiqlərində çox vaxt verilənlərin sonrakı emalı üçün sonda FFNN-dən istifadə edilir ki, bu da qeyri-xətti abstraksiyalar yaratmağa imkan verir.

6.5.3.2.5 Strukturlaşdırılmış perseptron

Strukturlaşdırılmış perseptron neyron şəbəkələri (NN) tərəfindən istifadə olunan

perseptron alqoritminin genişlənməsidir. Strukturlaşdırılmış perseptronlar çıxış verilənlərinin müxtəlif hissələrini proqnozlaşdırmaq üçün ardıcıl olaraq bir neyron layından bir neçə dəfə istifadə edir. Strukturlaşdırılmış perseptron təlimi proseduru bu prinsipi əks etdirir: bir məqsəd dəyişəni eyni zamanda bir neçə proqnozu (generasiya edilmiş nəticəni) qiymətləndirmək üçün istifadə olunur.

6.5.3.2.6 Dərin Bolsman maşını

Bolsman maşınları (BM) görünən vahidlər çoxluğu və gizli vahidlər çoxluğundan ibarət olan $\{0,1\}$ ikilik vahidlər şəbəkələridir. Əlaqələr yalnız qonşu vahidlər arasında var. BM ikiistiqamətli və naməlum ehtimal paylanmalarını öyrənməyə qadirdir. Onlar obyektlərin və ya nitqin tanınmasında faydalıdır. BM-lər generativ alqoritmdir.

BM-lərin ən sadə forması məhdud Bolsman maşınları (RBM) adlanır. RBM-in bir neçə layı ardıcıl (stek şəklində) yerləşdikdə, bu şəbəkəyə dərin Bolsman maşınları (DBM) deyilir. Dərin etimad şəbəkələri (DBN) DBM-ə bənzəyir, həm də neyron laylarını (yəni biristiqamətli əlaqələri) ehtiva edir.

6.5.3.2.7 Kapsul şəbəkələri

Kapsul neyron şəbəkələri (CapsNet) dinamik marşrutlaşdırmanı həyata keçirən və seyrək verilənlər üzərində öyrədilə bilən neyron şəbəkələridir. Onlar neyronları eyni vaxtda yenilənən və sıx əlaqələndirilmiş neyron çoxluqlarından ibarət olan “kapsul” adlanan strukturlarla əvəz etməklə, CNN-in bəzi məhdudiyyətlərini yaxşılaşdırır. CapsNet-in iyerarxik əlaqələri daha yaxşı əks etdirməsi və daha yüksək səviyyəli təqdimat generasiya edə bilməsi üçün kapsullar arasındakı əlaqələr də təkmilləşdirilir. CapsNet diskriminativ tipə aiddir.

6.5.3.2.8 Generativ rəqabətli şəbəkə

Generativ rəqabətli şəbəkələr (GAN) verilənlər çoxluğunu üçün ən reprezentativ nümunələri generasiya etməyə çalışan bir və ya bir neçə generatoru və generasiya olunan nümunələri real nümunələrdən fərqləndirməyə çalışan bir və ya bir neçə diskriminatoru ibarət neyron şəbəkələridir. Generatorun və diskriminatorun komponentləri şəbəkənin daxili təqdimatlarını təkmilləşdirmək üçün birlikdə öyrədilir.

GAN müxtəlif tapşırıqların (məsələn, klassifikasiya) yerinə yetirilməsi və həmçinin, süni verilənlərin yaradılması və ya yeni sahəyə adaptasiya kimi digər məqsədlər üçün istifadə edilə bilər.

6.5.3.3 Bayes şəbəkəsi

Bayes şəbəkələri dəyişənlər arasındakı asılılıqlar haqqında proqnozlar generasiya etmək üçün istifadə edilən qraf modelləridir. Onlar nəticəyə təsir edə biləcək səbəblərin və ya dəyişənlərin ehtimallarını müəyyən etmək üçün istifadə edilə bilər. Bu cür səbəb-nəticə əlaqəsi tibbi diaqnostika ilə bağlı tətbiqlərdə çox faydalıdır. Bayes şəbəkələri verilənlərin təhlili, natamam verilənlərlə iş və modellərin verilənlərə həddən artıq uyğunlaşmasının aradan qaldırılması üçün də faydalıdır. Bayes şəbəkələri Bayes ehtimalına əsaslanır: hadisənin ehtimalı həmin hadisənin baş verməsinə inamın dərəcəsi kimi başa düşülür. Səbəb-nəticə əlaqələrini müəyyən etmək və ya verilənlərin təhlilini aparmaq üçün Bayes

şəbəkələri ilə birgə istifadə edilə bilən müxtəlif Bayes statistik metodları mövcuddur. Bayes şəbəkələri tez-tez dəyişənlərin istiqamətləndirilmiş qeyri-dövri qraf adlanan qraf şəklində təqdimatından istifadə edir. Bu qrafların xüsusiyyəti ondan ibarətdir ki, x dəyişəni ilə digər dəyişənlər arasındakı əlaqələr üzrə keçdikdən sonra qraf heç vaxt x dəyişəninə qayıtmır. Bayes şəbəkələri haqqında ətraflı məlumat [8] istinadından əldə edilə bilər.

6.5.3.4 Naïve Bayes alqoritmi

Naïve Bayes teoreminə əsaslanan klassifikasiya metodudur. Teorem əvvəlki əlaqəli hadisələr barədə biliklərə əsaslanaraq, hadisənin baş vermə ehtimalını müəyyən edir. Bu biliklər hadisənin baş verib-verməyəcəyinin müəyyən edilməsinin dəqiqliyini artırmağa kömək edir. Məsələn, xəstənin oturaq və ya aktiv həyat tərzini nəzərə alaraq, xəstəliyin tibbi diaqnozu daha dəqiq qoyula bilər. Naïve Bayes klassifikatoru verilənlərin xüsusiyyətlərinin bir-birindən statistik asılı olmadığını nəzərdə tutur. Bu klassifikasiya metodunun üstünlükləri onun nisbətən sadə olması və təlim üçün böyük həcmdə verilənlər çoxluğunu tələb etməməsidir.

6.5.3.5 Dəstək vektoru maşını

Dəstək vektoru maşını (SVM) klassifikasiya və reqressiya üçün geniş istifadə olunan maşın öyrənməsi texnikasıdır. SVM klassifikatoru nümunələri işarələməklə iki fərqli kateqoriyaya təsnif edir və sonra yeni giriş verilənlərini bu kateqoriyalardan birinə təyin edir. SVM maksimum marjaya malik klassifikasiya alqoritmidir. SVM verilənləri iki sinfə ayırmaq üçün klassifikasiya müstəvisi ilə ən yaxın (verilənlər) nöqtələr arasında maksimum məsafəni təmin edən hiperüstəvi təyin edir. Sərhədə ən yaxın nöqtələr “dəstək vektorları” adlanır. Dəstək vektorları ilə hiperüstəvi arasında ortoqonal məsafə SVM “marja”sının yarısını təşkil edir. SVM, həmçinin verilənlərin giriş fəzasından (klassifikasiya edən) hiperüstəvinin seçiləcəyi daha yüksək ölçülü (bəzən sonsuz) fəzaya inikas olunması üçün nüvə funksiyalarından istifadə edir.

SVM təlimi hiperüstəvinin əks tərəflərindəki müxtəlif kateqoriyalı verilənlərə nəzərən marjanın maksimallaşdırılmasını ehtiva edir. Təcrübədə “sərt marjalı” (“hard-margin”) SVM-lər nadir hallarda istifadə olunur. “Sərt marjalı” SVM klassifikatoru yalnız verilənlərin daxili fəzada xətti olaraq klassifikasiya edilə bildiyi halda işləyir. Verilənlərin hətta bir nümunəsi bölgü aparıcı hiperüstəvinin tapılmasını qeyri-mümkün edə bilər. Bunun əksinə olaraq, “yumşaq marjalı” SVM klassifikatorları verilənlərin nümunələrinin marjanı pozmasına (yəni hiperüstəvinin digər tərəfində yerləşməsinə) imkan verir. “Yumşaq marjalı” SVM klassifikatorları marja pozuntularını məhdudlaşdırmaqla, maksimum marjaya nail olmağa çalışır. SVM tətbiqlərinə misal olaraq, nişanlanmamış verilənlərin kateqoriyalaşdırılmasını, eləcə də onların proqnozlaşdırma və obrazların tanınması məsələlərində istifadəsini göstərmək olar.

SVM-dən reqressiya məsələlərində istifadə edərkən, məqsəd SVM klassifikatorunun məqsədinin tam əksi olur. Belə ki, SVM reqressiyasında məqsəd marja pozulmalarını (yəni nümunələrin marjadan kənarında yerləşməsinə) məhdudlaşdırmaqla, mümkün qədər çox verilənlər nümunəsini marja daxilində yerləşdirməkdir.

6.5.3.6 Qərarlar ağacı

Qərarlar ağacı mümkün nəticələri kodlaşdırmaq məqsədilə qərarların ağacvari

strukturundan istifadə edir. Qərarlar ağacı alqoritmləri klassifikasiya və reqressiya məsələlərində geniş istifadə olunur. Ağac qərar təpələrindən və yarpaq təpələrindən ibarətdir. Hər bir qərar təpəsinin ən azı iki budağı var və burada yarpaqlar yekun qərarı və ya klassifikasiyanın təmsil edir. Adətən, təpələr ən güclü proqnozlaşdırıcı olan qərara görə sıralanır. Ən yaxşı nəticəni müəyyən etmək üçün giriş verilənləri müxtəlif amillər üzrə bölünməlidir. Qərarlar ağacı blok-sxemlərin analoqudur: burada hər bir qərar təpəsində növbəti hansı budağı seçəcəyinizi müəyyən etmək üçün sual verə bilərsiniz.

6.5.4 Maşın öyrənməsinin optimallaşdırma metodları

6.5.4.1 Ümumi müddəalar

ML optimallaşdırma metodları ML modelinin (parametrlərinin) ML verilənlərinə uyğunlaşdırılması üçün istifadə olunur. ML-in optimallaşdırılması məsələsi verilmiş ML verilənlər çoxluğu üçün verilmiş itki funksiyasını minimallaşdırmaq məqsədilə ML modelinin optimal parametrlərini müəyyən etməkdir. Bu metodların əksəriyyəti modelin məqbul həllə yığılmağını müəyyən etmək üçün itki funksiyasından (buna bəzən “xərclər” funksiyası da deyilir) istifadəyə əsaslanır. Geniş yayılmış ML optimallaşdırma metodları aşağıdakı altbəndlərdə təsvir edilmişdir.

Optimallaşdırma metodlarından başqa, digər metodlar da ML modelinin verilənlərə uyğunluğunu yaxşılaşdırma bilər. Belə aspektlərdən biri ən yaxşı itki funksiyasını seçməkdir. Məsələn, tənzimləmə - itki funksiyasına yeni hədlər daxil etməklə, həddən artıq öyrənməni və proqnozların yüksək dispersiyasını azaltmaq üçün istifadə edilən bir texnikadır. Tənzimləmə metodunun iş prinsipi modelin mürəkkəbliyinə görə cərimələr təyin etməkdən və daha ümumi, nisbətən az dəqiqliyə malik modelə üstünlük verməkdən ibarətdir.

6.5.4.2 Qradyent enişi metodları

Qradyent enişi funksiyanın minimumunu tapmaq üçün iterativ metoddur: bütün verilənlər çoxluğunun eyni zamanda vahid paket kimi təqdim olunduğu halda parametrlər funksiyanın birinci tərtib törəməsi (qradyenti) istiqamətində addım-addım yenilənir. Məqsədi funksiyası qabarıq olduqda, həll global optimumdur. Yenilənmiş parametrlərin bütün verilənlər çoxluğunda deyil, təsadüfi seçilmiş daha kiçik həcmdə verilənlər çoxluğunda hesablanması stoxastik qradyent enişi (SGD) adlanır. SGD hesablama xərclərini azaldır, lakin lokal minimuma uyğun həlli tapmaq riski var. Qradyentin iterativ addımının ölçüsü “öyrənmə sürəti” adlanır.

Moment (impuls) metodları (“momentum methods”) sürəti fəzada parametrlərin istiqamətini və dəyişmə sürətini təmsil edən dəyişən kimi daxil edir; bu əməliyyat növbəti iterasiya üçün əvvəlki yeniləmənin istiqamətinin təsirini saxlayır. Nesterovun sürətləndirilmiş qradyent enişi metodu adaptiv moment (impuls) anlayışını istifadə edir ki, bu da nəticədə daha sürətli yığılmanı təmin edir.

Hər bir parametr yeniləməsinin addım ölçüsünü dəyişdirərək öyrənmə sürətini tənzimləmək qradyent enişi metodlarını təkmilləşdirir. Buna “adaptiv öyrənmə sürəti” deyilir. Misal olaraq, öyrənmə sürətini avtomatik tənzimləyən AdaGrad (adaptiv qradyent alqoritmi) alqoritmini göstərmək olar.

Təlim verilənlərində izafi məlumatların mövcudluğu SGD metodlarında yavaş yığılmaya

səbəb ola bilər. Stoxastik orta qradiyent son qradiyentlərin cəmini qeyd etmək üçün parametrdən istifadə edir. Bu cəm hər iterasiyada təsadüfi olaraq yenilənir və hər dövrdə köhnə qradiyenti yeni qradiyentlə əvəz edir. Bu yanaşma dispersiyanın azaldılmasını təmin edir.

6.5.4.3 Nyuton metodu

Nyuton metodu məqsəd funksiyasının kvadratik funksiyalarla approksimasiyası üçün birinci tərtib törəmədən (qradiyent) və ikinci tərtib törəmə matrisindən (həmçinin Hesse matrisi kimi tanınır) istifadə edir. Nyuton metodu cari lokal səthi kvadratik səthə uyğunlaşdırır; bu da Nyuton metodunu cari lokal səthi müstəviyə uyğunlaşdıran qradiyent enişi metodundan fərqləndirir.

6.5.4.4 Qoşma qradiyent

Qoşma qradiyent (CG) metodu Nyuton metodu kimi Hesse matrisini birbaşa hesablamadan Hesse vektor hasilini hesablayır. CG yalnız əvvəlki vektordan istifadə edərək yeni axtarış istiqaməti generasiya edir. Bu, tərs Hesse matrisi ilə bağlı hesablama xərclərinin qarşısını alır.

6.5.4.5 Gauss prosesləri

Gauss prosesi (GP) ardıcıl ümumi Gauss paylamalarına malik təsadüfi dəyişənlər toplusudur. GP Bayes nəzəriyyəsinə və statistik öyrənməyə əsaslanır. Metodun üstünlüklərinə çevik qeyri-parametrik nəticə çıxarma, eləcə də istifadə mürəkkəbliyi və böyük həcmli yaddaş tələblərinə baxmayaraq, yüksək uyğunluq ("interoperability") daxildir.

6.5.4.6 Ən kiçik kvadratlar metodu ilə approksimasiya

Ən kiçik kvadratlar metodu ilə approksimasiya - polinomial (çoxhədli) funksiyanın qiymətləri ilə verilənlər arasındakı fərqlərin kvadratlarının cəmini minimallaşdırmaqla, polinomial funksiyanın verilənlərə uyğunlaşdırılması (yaxınlaşdırılması) üsuludur. Yaxşı nəticələr əldə etmək üçün çoxhədlinin tərtibi ölçmə verilənlərinə ("measurement data") uyğun seçilməlidir. Ən kiçik kvadratlar metodu ilə approksimasiya tam verilənlər çoxluğu üzərində paket şəklində və ya artan verilənlər çoxluğu üzərində rekursiv şəkildə həyata keçirilə bilər.

6.5.4.7 Maksimum ehtimalın qiymətləndirilməsi

Maksimum ehtimalın qiymətləndirilməsi ehtimal funksiyasının maksimallaşdırılması ilə ehtimalın paylanması parametrlərinin qiymətləndirilməsi metodudur. Müşahidə olunan verilənlər əsasında ehtimal modelini qurmaq üçün maksimum ehtimalın qiymətləndirilməsi metodunda müşahidə edilən verilənlərin ehtimalını maksimallaşdıran hipotez (fərziyyə) seçilir. Ümumi halda bu metod hər hansı konkret hipotezə üstünlük vermək üçün apriori biliklərdən istifadə etmir. Maksimum ehtimalın qiymətləndirilməsi metodunda hesablamaların aparılması üçün çox vaxt ehtimalın loqarifmindən istifadə edilir.

6.5.4.8 Gözləmənin maksimallaşdırılması

Gözləmənin maksimallaşdırılması (EM) verilənlərdə açıq şəkildə müşahidə olunmayan

gizli dəyişənlərə malik model parametrlərini müəyyənləşdirmək üçün istifadə olunan iterativ metoddur. Bu metod yığılma əldə olunana qədər gözləmə addımlarının (cari parametrlər əsasında gizli dəyişənlərin qiymətləndirilməsi) və maksimallaşdırma addımlarının (gizli dəyişənlərin cari qiymətini nəzərə alaraq verilənlərin ehtimalının optimallaşdırılması üçün parametrlərin yenidən qiymətləndirilməsi) növbələşməsindən ibarətdir.

6.5.5 Maşın öyrənməsinin qiymətləndirilməsi metrikaları

6.5.5.1 Ümumi müddəalar

ML tapşırıqları (6.2-ci bənddə müzakirə olunur) adətən müvafiq qiymətləndirmə metrikalarını müəyyən edir. Modelin istifadə olunacaq variantını “başa düşmək” müvafiq qiymətləndirmə metrikasını seçmək üçün vacibdir. Modelin performansını adekvat şəkildə ifadə etmək və araşdırmaq üçün bir neçə metrika tələb oluna bilər. Əsas nəticələrin hərtərəfli yoxlanılmasını aparmadan bir metrikadan, hətta F1 kimi aqreqsiya olunmuş bir metrikadan istifadə etmək risklidir. Bundan əlavə, təlim verilənlərində siniflərin balanslaşdırılmaması müxtəlif metrikaları təhrif edə bilən amildir.

Bu metrikaların bir çoxu (xüsusən də, klassifikasiya məsələlərində) doğru müsbət (T_P , düzgün aşkar edilən nümunələr), yanlış müsbət (F_P , aşkar edilməməli olduğu halda səhvən aşkar olunan nümunələr), doğru mənfi (T_N , aşkar edilməməli olan nümunələrin düzgün müəyyən olunması) və yanlış mənfi (F_N , aşkar edilməli olduğu halda səhvən aşkar edilməyən nümunələr) anlayışlarına əsaslanır.

Təlim keçmiş modellərin performansını əks etdirmək üçün bir çox metrikalar mövcuddur, məsələn:

— Klassifikasiya alqoritmlərini qiymətləndirmək üçün doğruluq (“accuracy”), qəbuledicinin iş xüsusiyyətləri (ROC), xəta matrisi, dəqiqlik (“precision”), dolğunluq (“recall”) və F1 meyarından istifadə etmək olar;

— Orta mütləq səhv (MAE), orta kvadratik səhv, nisbi mütləq səhv, nisbi kvadratik səhv, orta “sıfır-bir” səhvi və determinasiya əmsalı reqressiya modelləri üçün ümumi metrikalardır;

— Klasterizasiya modelləri üçün qiymətləndirmə metrikalarına klasterin mərkəzinə orta məsafə, digər mərkəzə olan orta məsafə, nöqtələrin sayı və klasterin mərkəzinə maksimal məsafə daxildir.

Aşağıdakı altbəndlər bu metrikalara dair bəzi nümunələri təsvir edir.

6.5.5.2 Dəqiqlik, dolğunluq, həssaslıq və spesifikli

Bir sıra metrikalar birbaşa T_P , F_P , T_N və F_N arasındakı nisbətələr kimi hesablanı bilər:

— Doğru müsbət hallar tezliyi (TPR) doğru müsbət halların (T_P) bütün faktiki müsbət hallara ($T_P + F_N$) nisbətidir. TPR-in tamamlayıcısı yanlış mənfi hallar tezliyi (FNR) adlanır;

— Doğru mənfi hallar tezliyi (TNR) doğru mənfi halların (T_N) bütün faktiki neqativ hallara ($T_N + F_P$) nisbətidir. Onun tamamlayıcısı yanlış müsbət hallar tezliyi (FPR) adlanır;

— Müsbət proqnoz ehtimalı (PPV) doğru müsbət halların (T_P) bütün proqnozlaşdırılan müsbət hallara ($T_P + F_P$) nisbətidir.

— Mənfi proqnoz ehtimalı (NPV) doğru mənfi halların (T_N) bütün proqnozlaşdırılan mənfi hallara ($T_N + F_N$) nisbətidir.

Bir-birini tamamlayan baxış bucaqları təklif etdikləri üçün bu metrikalar qiymətləndirmədə cüt istifadə olunur: ya PPV və TPR, həmçinin dəqiqlik və dolğunluq adlanır; ya da TPR və TNR, həmçinin həssaslıq və spesifiklik adlanır.

Bu tərifləri tətbiq etməyin bir çox yolları var, terminlərin mənası isə klassifikasiyanın binar və ya çoxsınıflı olmasından asılı olaraq fərqlənir. Binar klassifikasiyada bir sinif müsbət olaraq seçilir və təriflər yalnız bu sinfə tətbiq olunur. Beləliklə, məsələn, dəqiqlik termini müsbət sinfin dəqiqliyinə aid olur və digər sinfin dəqiqliyi hesablanmır. Çoxsınıflı klassifikasiyada təriflər öz növbəsində bütün siniflərə tətbiq edilir, dəqiqlik isə hər sinfin dəqiqliyinin hər hansı kombinasiyasını ifadə edir.

6.5.5.3 F1 meyarı

F1 meyarı dolğunluğun və dəqiqliyin kombinasiyası vasitəsilə modelin performansını ifadə edir. F1 meyarı dəqiqlik və dolğunluğun orta harmonik qiyməti ilə müəyyən edilir. F1 meyarının ən yaxşı qiyməti 1-ə bərabərdir və bu, mükəmməl dəqiqliyin və dolğunluğun göstəricisidir.

6.5.5.4 Doğruluq

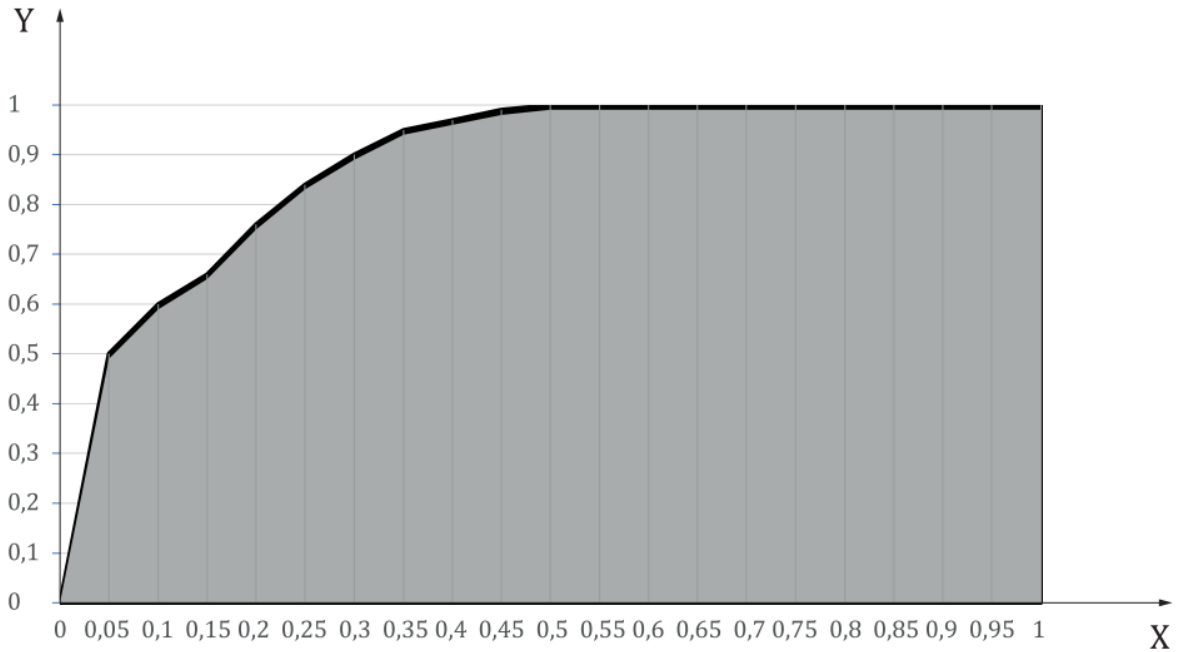
Doğruluq modelin ümumi yanlış və düzgün klassifikasiyaları arasında düzgün klassifikasiyaların payını (doğruları) ifadə edir.

Əsasən, binar klassifikasiya hallarında aktualdır. Çoxsınıflı klassifikasiyada bunu ya hər bir sinif üçün doğruluq meyarı, ya da modelin doğruluq meyarı kimi başa düşmək olar. Hər sinif üzrə doğruluq həmin sinfin dolğunluğuna bərabər götürülür və “dolğunluq” termininə üstünlük verilir. Modelin doğruluğu nümunənin aid olduğu konkret sinfi nəzərə almır və yalnız bütün nümunələrin nəticələrini fərq qoymadan cəmləməklə, onun klassifikasiyasının düzgün olub-olmamasını nəzərə alır. Lakin bu metrika adətən dəqiqlik və dolğunluq cütüyündən və ya F1 meyarından daha az informativ olur.

Çoxqiymətli klassifikasiyada “doğruluq” termini çox vaxt altçoxlğun doğruluğuna aid olur, yəni düzgün nişanlar (“label”) çoxluğunun (bütövlükdə) proqnozlaşdırılmasının doğruluğunu ifadə edir.

6.5.5.5 Qəbuledicinin iş xüsusiyyətləri və əyri altındakı sahə

ROC (Qəbuledicinin iş xüsusiyyətləri) modelin sinifləri fərqləndirmək qabiliyyətini əks etdirən əyri şəklində modelləşdirilir və TPR ilə FPR arasındakı asılılığı göstərir. Əyri üzərində istənilən nöqtə model üçün TPR və FPR haqqında məlumatları təmin edən əməliyyat (operating) nöqtəsi kimi seçilə bilər. Beləliklə, tətbiq sahəsindən və TPR ilə FPR göstəricilərinə verilən əhəmiyyətdən asılı olaraq, arzuolunan dəqiqliyi və dolğunluğu maksimallaşdırmaq üçün əməliyyat nöqtəsi seçilə bilər. Şəkil 4-də ROC əyrisinin nümunəsi təqdim edilir.



X - yanlış müsbət hallar tezliyi, FPR

Y - doğru mənfi hallar tezliyi, TPR

■ - AUC sahəsi

■ - ROC əyrisi

Şəkil 4 — Qəbuledicinin iş xüsusiyyətləri əyrisi

ROC əyrisi altında olan sahə (AUC) klassifikasiyanın bütün sərhəd qiymətləri üzrə performans (effektivlik) ölçüsüdür. AUC-un qiyməti həmişə $[0, 1]$ diapazonundadır və modelləri rəqləmə üçün istifadə edilə bilər. AUC-un daha yüksək qiymətinə proqnozlaşdırmanın daha yaxşı nəticəsi uyğundur. AUC-un mümkün ən yaxşı qiyməti 1,0, gözlənilən ən pis qiyməti isə 0,5-dir (çünki diaqonal xətt təsadüfi proqnozlara uyğundur). 0,5-dən az olan istənilən qiyməti o deməkdir ki, istifadəçi qiymətin 0,5-dən yuxarı olması üçün modelin tövsiyə etdiyi addımların tam əksini ata bilər [9]. AUC miqyasa və klassifikasiyanın sərhəd qiymətlərinə nəzərən invariantdır və bu, tətbiqindən asılı olaraq, həm yaxşı, həm də mənfi xüsusiyyət ola bilər. Adətən təlim verilənlərində siniflərin balansız pozulan zaman AUC-dan istifadə olunur. Şəkil 4-də AUC nümunəsi ROC əyrisinin altındakı boz sahə kimi göstərilir.

6.5.5.6 Xətalər matrisi

Xətalər matrisi ("confusion matrix" - bəzən "qarışıqlıq matrisi" də adlanır) klassifikasiya modelinin effektivliyini vizuallaşdırmağa imkan verən cədvəldir. Adətən sətirlər müəyyən bir sinfə aid olması proqnozlaşdırılan elementlərin sayını, sütunlar isə həqiqətən aid olduqları siniflərindəki (yəni düzgün klassifikasiya olunmuş) elementlərin sayını göstərir. Diaqonalda nə qədər çox element cəmləşsə, siniflər arasında qarışıqlıq bir o qədər az, modelin performansı isə bir o qədər yaxşı olar. Şəkil 5-də iki sinfə bölən model əks olunur. Bu konkret binar klassifikasiyada yuxarı sol xana doğru müsbət halların (T_P), yuxarı sağ xana yanlış müsbət halların sayını (F_P), aşağı sol xana yanlış mənfi halların sayını (F_N) və

aşağı sağ xana isə doğru mənfi halların sayını (T_N) ehtiva edir. Xətlər matrisi TPR, FPR, Dəqiqlik, F1 meyarı və digər metrikaları hesablamağa imkan verir.

		Real siniflər	
		Sinif 0	Sinif 1
Proqnozlaşdırılan siniflər	Sinif 0	T_P	F_P
	Sinif 1	F_N	T_N

Şəkil 5 — Xətlər matrisi

Klassifikasiya ilə yanaşı, xətlər matrisini bir sıra digər tapşırıqlar üçün də tərtib etmək mümkündür. Bununla belə, tapşırıqdan asılı olaraq, cədvəl qismən doldurula bilər: məsələn, obyektlərin aşkarlanması tapşırığında T_N qiyməti çətin müəyyən edilir.

6.5.5.7 Kappa əmsalı

Kappa əmsalı (həmçinin “Cohen kappa əmsalı” kimi də tanınır) kateqorial elementlər üçün qiymətləndiricilər (“rater”) arasındakı etibarlılığı ölçür. Maşın öyrənməsində kappa əmsalı modelin klassifikasiyası ilə məlumatların etalon verilənlərin nişanları arasındakı uyğunluğu ölçmək üçün istifadə olunur. O, həmçinin bir neçə modelin uyğunluğunu yoxlamaq və etibarlılığı aşağı olan modellərdən imtina etmək üçün də istifadə edilə bilər. Kappa əmsalı balanslaşdırılmayan verilənlər çoxluqları üçün də faydalıdır, belə ki, bu əmsal tətbiqin hədəf paylanmasıdan istifadə edərək, modelin performansını təsadüfi təxminlərə nəzərən ifadə edir.

(k) $\{0,1\}$ əmsalı müşahidəçilər arasındakı razılıq dərəcəsini ölçür. Məlumat toplayanların (qiymətləndiricilər) eyni dəyişənə və ya müşahidə olunan obyektə eyni qiymət və ya bal təyin etməsinin ölçülməsinə “qiymətləndiricilər arasındakı etibarlılıq” deyilir. Bu ölçü subyektiv qiymətləndiricilərin qeyri-müəyyənlik səbəbindən dəyişənlərin qiymətlərini təxmin etmə ehtimalını əks etdirir.

Bəzən model performansını qiymətləndirmək üçün bu metrikadan istifadə olunsada, o bu məqsədlər üçün nəzərdə tutulmayıb və elmi ədəbiyyatda bu cür istifadə tövsiyə edilmir [10] [11].

6.5.5.8 Metyu korrelyasiya əmsalı

Metyu (Matthew) korrelyasiya əmsalı (MCC) binar klassifikasiyanın keyfiyyətini əks etdirir. Statistik olaraq “fi” (ϕ) əmsalı kimi tanınır və xi-kvadrat (χ^2) statistikasına əlaqəlidir. Əmsal bilavasitə xətlər matrisi vasitəsilə hesablanır (bax: 6.5.5.6):

$$\varphi = \frac{T_P \times T_N - F_P \times F_N}{\sqrt{(T_P \times F_P)(T_P \times F_N)(T_N \times F_P)(T_N \times F_N)}}$$

7 MAŞIN ÖYRƏNMƏSİ YANAŞMALARI

7.1 Ümumi müddəlar

Maşın öyrənməsi üsulları üç yanaşma üzrə klassifikasiya edilə bilər: müəllimlə maşın öyrənməsi, müəllimsiz (supervayzersiz) maşın öyrənməsi və möhkəmləndirici öyrənmə. Qismən müəllimlə maşın öyrənməsi, öz-özünə (sərbəst) öyrənmə, transfer öyrənmə və ansambl öyrənməsi eyni zamanda bir neçə ML yanaşmasına əsaslanır və onlar ayrıca müzakirə mövzudur.

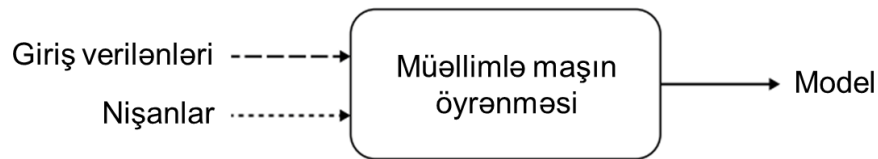
Növbəti bölmələrdə təsvir edilən ML yanaşmaları Şəkil 6-da göstərildiyi kimi müxtəlif ML metodlarından istifadə edə bilər. İstifadə olunan metodların bəziləri 6-cı bölmədə təsvir edilmişdir. Şəkil 6-da təqdim olunan siyahılar nümunə xarakteri daşıyır və bütün metodlar bu standartda təsvir olunmur. Şəkil 6 bəzi metodların bir neçə ML yanaşmasında istifadə edildiyini göstərir. Məsələn, neyron şəbəkələri ML bütün yanaşmalarında istifadə olunur.

Müəllimlə maşın öyrənməsi		Müəllimsiz maşın öyrənməsi	
Klassifikasiya	Reqrəssiya	Klasterizasiya	Ölçülərin azaldılması
<ul style="list-style-type: none"> • Logistik reqrəssiya • Xətti diskriminant analizi • Naive Bayes • K-ən yaxın qonşu metodu • Qərarlar ağacı • Ansambl metodu (təsadüfi meşə, AdaBoost) • Nüvə metodu, dəstək vektoru metodu • Neyron şəbəkələri 	<ul style="list-style-type: none"> • Xətti və qeyri-xətti əyrilərlə yaxınlaşma • Reqressiya ağacı • Ridge və Lasso reqrəssiya • Bayes reqrəssiya • ElasticNet reqrəssiya • Dəstək vektoru reqrəssiya • Qaus prosesi reqrəssiya • Neyron şəbəkələri 	<ul style="list-style-type: none"> • K-orta klasterizasiya • İyerarxik klasterizasiya • Qaus qarışığı modeli • Sıxlığa əsaslanan klasterizasiya • Genetik alqoritmlər • Spektral klasterizasiya • Miqyas-fəza klasterizasiya • Neyron şəbəkələri 	<ul style="list-style-type: none"> • Əsas komponentlər metodu • Tenzor ayrılışı • Xüsusiyyətlərin izometrik inikası • Lokal xətti quraşdırılma • T-paylanma ilə qonşuların stoxastik quraşdırılması • Müntəzəm approksimasiya və proyeksiya ("Uniform manifold approximation and projection") • Neyron şəbəkələri
Möhkəmləndirici maşın öyrənməsi			
Modelsiz		Model əsaslı	
<ul style="list-style-type: none"> • Monte Karlo • Mövcud vəziyyət/mövcud fəaliyyət/növbəti mükafat/növbəti vəziyyət/növbəti fəaliyyət • Q-öyrənmə, dərin Q-öyrənmə • Trayektoriya üzrə inteqral ilə siyasətin təkmilləşdirilməsi • Etimad bölgəsi siyasətinin optimallaşdırılması, TRPO • Dərin determinik siyasət qradiyenti, DDPG • Proksimal siyasətin optimizasiyası, PPO • Neyron şəbəkələri 		<ul style="list-style-type: none"> • Dinamik proqramlaşdırma • Aşkar tədqiqat və istismar ("Explicit explore/exploit") • Model əsaslı interval qiymətləndirilməsi • Bayes əsaslı interval qiymətləndirilməsi • Differensial dinamik proqramlaşdırma • Neyron şəbəkələri 	

Şəkil 6 — Müəllimlə maşın öyrənməsi, müəllimsiz maşın öyrənməsi və möhkəmləndirici öyrənmə üzrə klassifikasiya olunan ML metodları

7.2 Müəllimlə maşın öyrənməsi

Müəllimlə (supervayzerlə) maşın öyrənməsində ML modelləri nişanlanmış verilənlərdən istifadə etməklə öyrədilir. Nişanlanmış verilənlər girişləri düzgün və ya doğru çıxış verilənlərinə uyğunlaşdırılmış nümunələrdən ibarətdir. Beləliklə, təlim verilənləri giriş dəyişənləri və “doğru” çıxış verilənlərindən ibarət cütlük təşkil edilir. Müxtəlif kontekstlərdə doğru çıxış verilənlərinə nişanlar, məqsəd dəyişənləri və etalon verilənlər də deyildir. Şəkil 7-də göstərilədiyi kimi, müəllimlə öyrənmə prosesi zamanı model yaratmaq üçün alqoritm giriş və çıxış verilənlərinə uyğunlaşdırılır.

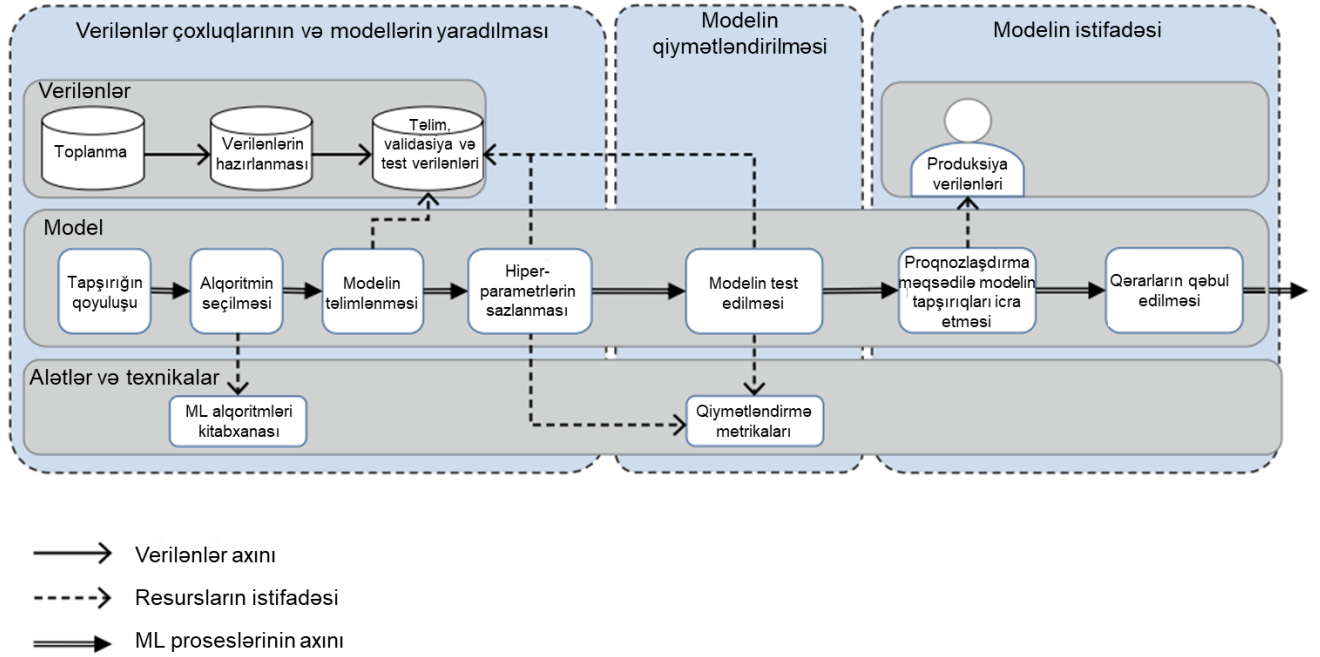


Şəkil 7 — Müəllimlə maşın öyrənməsindən istifadə edərək ML modelinin qurulması

Nişanlar ilkin verilənlərin bir hissəsi ola bilər, lakin əksər hallarda nişanları əl ilə və ya digər Sİ prosesləri (texnologiyaları) vasitəsilə yaratmaq lazım olur. Hədəf ML tapşırığından asılı olaraq, nişanlar müxtəlif formalarda ola bilər:

- klassifikasiya kateqorial nişanlar tələb edir (məsələn, it və ya bina kimi verilənlər nümunəsinin aid olduğu kateqoriya);
- reqressiya üçün nişanların ədədi olması tələb olunur (dərəcələr, oxşarlıqlar və ya ehtimallar kimi kəsilməz qiymətlər);
- strukturlaşdırılmış proqnozlaşdırma ilə bağlı onlar strukturlaşdırılmış obyekt (ardıcılıq, şəkil, ağac və ya qraf kimi) formasını da ala bilər.

ML modelinin qurulması, qiymətləndirilməsi və istifadəsi zamanı müxtəlif prosesləri əks etdirən Şəkil 8-də müəllimlə maşın öyrənməsi prosesinin tipik nümunəsi göstərilir. “Verilənlər çoxluğunun formalaşdırılması və modelin qurulması” mərhələsi modelin hazırlanması, təlimi və seçilməsi, eləcə də modelin qurulması və ya qiymətləndirilməsi üçün lazım olan verilənlərə uyğundur. “Modelin qiymətləndirilməsi” mərhələsində onun performansını və uyğunluğunu qiymətləndirmək üçün qiymətləndirmə metrikalarından istifadə etməklə model test edilir. “Modeldən istifadə” mərhələsində model proqnozlar vermək üçün produksiya verilənləri üzərində tətbiq edilir. Şəkil 8-də üfüqi sahələr üç mərhələdən ibarət olsa da, şaquli sahələr təsvir olunan komponent və proseslərin verilənlər, model və ya alətlərlə əlaqəli olub-olmadığını göstərir.



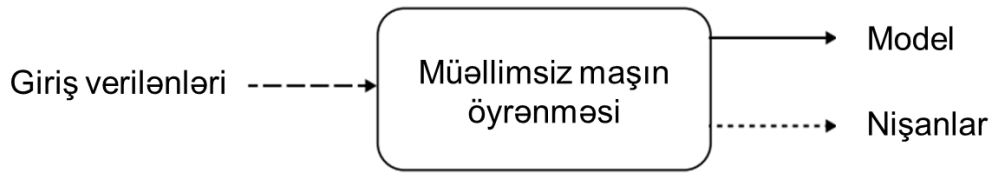
Şəkil 8 - Tipik müəllimlə maşın öyrənməsi prosesi

Təlim keçmiş modelin performansını və dayanıqlığını daha çox təlim verilənlərinin müxtəlifliyinə (məsələn, müxtəlif kateqoriya piyadalar), təlim verilənlərinin keyfiyyətinə (məsələn, fotoların işıqlandırılma effektləri və ya ayırdetmə dəqiqliyi) və nişanlama dəqiqliyinə (məsələn, piyada keçidində piyadanın düzgün nişanlanması) əsaslanır. Müəllimlə maşın öyrənməsi prosesindəki verilənlərin bütün aspektləri xəyata məruz qala bilər və ona görə də verilənlər çoxluğunun yaradılmasından modelin test edilməsinə qədər tam dövr ərzində bu prosese xüsusi diqqət yetirilməlidir.

7.3 Müəllimsiz maşın öyrənməsi

Müəllimlə maşın öyrənməsindən fərqli olaraq, müəllimsiz (supervayzersiz) maşın öyrənməsi (nişanlanmış verilənlər üzərində təlim keçmədən) giriş verilənlərini birbaşa çıxış verilənlərinə inikas edir. Bununla belə, təlim prosesi Şəkil 8-də əks olunan müəllimlə maşın öyrənməsinə bənzəyir. Şəkil 9-da verilən müəllimsiz maşın öyrənməsində alqoritm (öncədən nişanlara müraciət etmədən) model yaratmaq üçün yalnız giriş verilənlərinə uyğunlaşdırılır. Nişanlar adətən modelin təlimlənmə prosesinin kənar elementi kimi yaranır.

Alqoritmlərdən (məsələn, k -orta alqoritm) istifadəni ehtiva edən klasterləşmə tapşırıqlarında nümunələr hər bir nümunədən mərkəzə (sentroidə) qədər olan minimum məsafə əldə edilənə qədər iterativ olaraq klasterləşdirmə alqoritmii ilə emal olunur. Ölçülərin azaldılması tapşırıqlarında (məsələn, əsas komponent təhlili (PCA) alqoritmindən istifadə etməklə) giriş verilənlərinin hər bir xüsusiyyətinin dispersiyası hesablanır və əvvəlcədən müəyyən edilmiş sayda ən böyük dispersiyaya malik xüsusiyyətlər qaytarılır. Bu şəkildə işlənmiş modellər oxşarlıqları, qanunauyğunluqları (obrazları) və ya anomaliyaları, həmçinin ölçülərin azaldılmasını (burada statistik ən mühüm xüsusiyyətlər nişanlardan asılı olmayaraq təyin olunur) müəyyən etmək üçün istifadə edilə bilər. Müəllimsiz maşın öyrənməsi çox vaxt yeni biliklərin aşkarlanmasına gətirib çıxarır.



Şəkil 9 — Müəllimsiz maşın öyrənməsindən istifadə edərək ML modelinin qurulması

Müəllimsiz maşın öyrənməsindən aşağıdakı hallarda istifadə edilir:

- giriş verilənləri çoxluğunda klasterlərin aşkar edilməsi;
- latent amillərin aşkar edilməsi. Yüksək ölçülü verilənlər üçün çox vaxt onların “mahiyətini” əks etdirən daha aşağı ölçülü altfəzaya proyeksiya etməklə, ölçüsünün azaldılması arzuolunandır;
- bir neçə dəyişənin ölçülər çoxluğunda korrelyasiyaların müəyyən edilməsi;
- zədələnmiş və ya digər qüsurları olan şəkillərin bərpa (retuş) edilməsi;
- adətən, birlikdə alınan və ya satılan məhsulların qruplarının müəyyən edildiyi bazar (istehlak) səbətinin təhlili.

7.4 Qismən müəllimlə maşın öyrənməsi

Qismən müəllimlə maşın öyrənməsi “təlim zamanı həm nişanlanmış, həm də nişanlanmamış verilənlərdən istifadə edən maşın öyrənməsi” kimi müəyyən edilir. Qismən müəllimlə maşın öyrənməsi müəllimlə və müəllimsiz maşın öyrənməsinin hibrid variantıdır.

Qismən müəllimlə maşın öyrənməsinə digər bir yanaşma isə modelin ümumi performansını artırmaq üçün psevdo-nişanların yaradılması və istifadəsidir. Bu halda model ilkin olaraq nişanlanmış verilənlərdən istifadə etməklə təlimlənilir. Daha sonra təlim keçmiş modeldən nişanlanmamış verilənlərin nümunələri üçün psevdo-nişanları proqnozlaşdırmaq üçün istifadə olunur. Nəticədə təlim verilənləri çoxluğu həm nişanlanmış, həm də psevdo-nişanlanmış nümunələrdən istifadə edilərək tərtib edilir və modelə yenidən təlim keçmək üçün istifadə olunur. Bu yanaşma “özünütəlim” (“sərbəst təlim”) adlanır.

Qismən müəllimlə maşın öyrənməsi böyük təlim verilənləri çoxluğunda bütün nümunələrin nişanlanması vaxt və ya xərc nöqtəyi-nəzərindən mümkün olmadığı hallarda faydalıdır.

7.5 Öz-özünə maşın öyrənməsi

Öz-özünə (sərbəst) maşın öyrənməsi adətən müəllimlə maşın öyrənməsinə aid olan alqoritmlərdən istifadə etməklə nişanlanmamış verilənlər üzərində təlim keçmək üçün bir yanaşmadır. Bu yanaşma bilavasitə giriş verilənləri (məsələn, oxşarının yaradılması üçün təsvir), giriş verilənlərinin bir hissəsi (məsələn, giriş cümləsindəki bir söz) kimi qeyri-aşkar nişanlardan və ya “çiy” verilənlərdən (məsələn, süni şəkildə qarışdırılmamış ardıcılığın

qarışdırılmamış versiyası) asanlıqla generasiya edilə bilən hər hansı digər nişanlardan istifadə etməklə həyata keçirilir. Bir qayda olaraq, o, böyük həcmli verilənlərdə tətbiq olunur.

Öz-özünə maşın öyrənməsi ən çox təqdimatları öyrətmək üçün istifadə olunur: modelin son çıxış verilənləri silinir, aralıq çıxış verilənləri isə başqa maşın öyrənməsi modelində giriş kimi istifadə edilə bilər. Bu, xüsusiyyətlərin müəyyənləşdirilməsi üzrə səyləri (xərcləri) azaltmaqla yanaşı strukturlaşdırılmamış verilənlərin emalı üçün də xüsusən faydalıdır.

Nəzərə alınmalıdır ki, öz-özünə maşın öyrənməsi qismən müəllimlə maşın öyrənməsinin xüsusi metodu olan öz-özünə təlim deyil.

7.6 Möhkəmləndirici maşın öyrənməsi

Möhkəmləndirici maşın öyrənməsi digər yanaşmalardan fərqlənir, çünki onun əsas prinsipi modelin müəyyən vəziyyətdə inisializasiyasından, fəaliyyətin ("action") yerinə yetirilməsindən, həmin fəaliyyətə görə mükafatın müəyyən edilməsindən və modelin mükafatı maksimallaşdırmağa çalışan yeni vəziyyətə keçməsinə ibarətdir. Təlim modelin inisializasiyası və ya modelin fəaliyyətləri həyata keçirmək üçün istifadə etdiyi siyasəti müəyyən etmək məqsədilə istifadə oluna bilər.

Möhkəmləndirici maşın öyrənməsi əvvəlcədən müəyyən edilmiş məqsədə çatmaq üçün bir və ya bir neçə agentin öz mühiti ilə qarşılıqlı əlaqəyə öyrədilməsi prosesidir. Gücləndirici öyrənmədə maşın öyrənməsi agentləri "cəhd et və səhv et" iterativ prosesi vasitəsilə öyrənirlər. Agentin məqsədi öz ətrafında ən yaxşı mükafatları təmin edən strategiya tapmaqdır (yəni model qurmaqdır). Hər bir (uğurlu və ya uğursuz) cəhd üçün mühit dolayı əks əlaqə təmin edir. Daha sonra agent bu əks əlaqə rəyi əsasında davranışını (yəni modeli) tənzimləyir. Bu proses Şəkil 10-da təsvir edilmişdir. Agent məqsədinə nail olmaq üzrə fəaliyyətində ardıcıl olaraq hansı qarşılıqlı əlaqələrin maksimum mükafat təmin etdiyini müəyyən edir.

Agentin fəaliyyəti və ətrafla qarşılıqlı əlaqəsi adətən Markov qərar qəbuletmə prosesi (MDP) kimi modelləşdirilir. Şəkil 10 MDP-nin tipik təsviridir. Möhkəmləndirici maşın öyrənməsi alqoritmləri MDP-nin dəqiq riyazi modeli haqqında məlumatı nəzərdə tutmur (lakin bəzi üsullar onu approksimasiya etməyə çalışır). Möhkəmləndirici maşını öyrənməsi alqoritmləri adətən dəqiq metodların tətbiqinin mümkün olmadığı böyük MDP modellərini hədəfə alır. Müəllimlə maşın öyrənməsindən fərqli olaraq, giriş verilənləri və onlarla əlaqələndirilmiş nişanlanmış doğru çıxış verilənləri cütü tələb olunmur. Bunun əvəzinə, məqsəd "cəhd et və səhv et" üsulundan istifadə etmək və "təxirə salınmış" mükafatlar vasitəsilə nəticələri müəyyən olunmuş məqsədə çatdırmaqdır. Model hər dəfə proqnoz verəndə mükafat hesablanır və mükafatı optimallaşdırmaq üçün əlavə cəhdlər edilir. Mükafat adətən sistemin konkret cəhd çərçivəsində məqsədə çatmağa nə qədər yaxın olduğunu əks etdirən (hesablanmış) ədəddir. Bəzi hallarda, təlimin nəticələrini yaxşılaşdırmaq üçün ətraf mühit haqqında əlavə məlumat modelləşdirilir və ya agentə əlavə məlumat verilir. Məqsəd və ya müvəffəqiyyətin əldə edilməsi (tərfi) adətən sistemin layihələndiricisi tərəfindən müəyyən edilir.

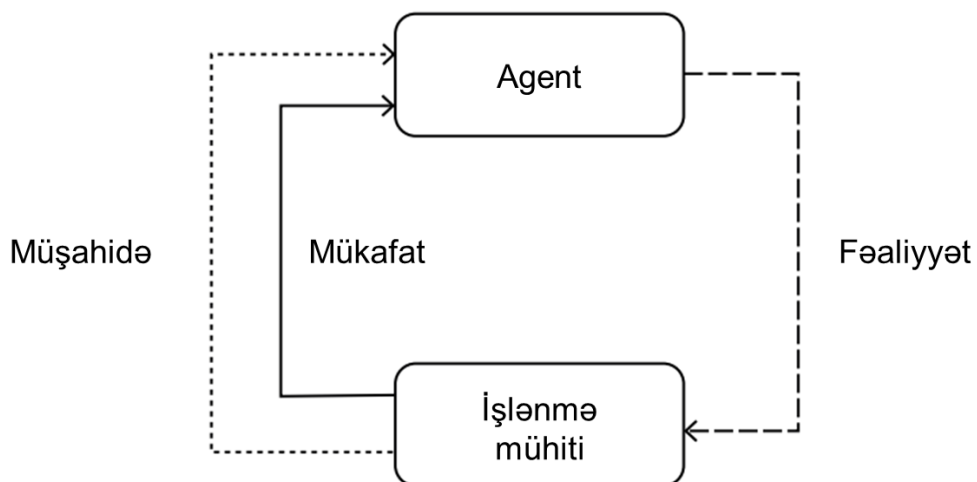
Möhkəmləndirici maşın öyrənməsi müəllimlə öyrənmə ilə birgə də istifadə edilə bilər: nişanlanmış verilənlər üzrə təlim modelin inisializasiyası üçün, möhkəmləndirmə isə agentin fəaliyyətlərini həyata keçirməsi məqsədilə istifadə etdiyi sonrakı strategiyaları müəyyən etməsi üçün istifadə olunur.

Bəzi hallarda (məsələn, strukturlaşdırılmış proqnozlaşdırmada) möhkəmləndirici maşın öyrənməsinin tətbiqi mühit yaratmaq üçün tələb olunan əvvəlcədən mövcud (nişanlanmış) verilənlər çoxluğundan asılı ola bilər: belə hallarda mühit verilənlər çoxluğuna daxil edilmiş məlumatlar üçün vasitəçi (proksi) rolunda çıxış edir.

Möhkəmləndirici maşın öyrənməsi fasiləsiz öyrənmənin tətbiqini asanlaşdırır, çünki istismar mərhələsində mükafatı hesablamaq üçün kifayət qədər məlumatın olması şərti ilə (bu mərhələdə) təlim mühiti simulyasiya edilmiş və ya real mühit ola bilər.

Möhkəmləndirici maşın öyrənməsi tez-tez idarəetmə məqsədləri üçün istifadə olunur. İdarəetmə - ətraf mühitlə qarşılıqlı əlaqə üçün gücləndirici maşın öyrənməsinin tətbiqidir. Mükafatı maksimallaşdıran siyasətin (qaydaların) müəyyən edilməsi fəaliyyətin planlaşdırılmasına və optimal idarəetməyə imkan yaradır, beləki fəaliyyətlər bu siyasətə uyğun həyata keçirilir.

Möhkəmləndirici maşın öyrənməsindən istifadə hallarına stolüstü və video-oyunlar daxildir, burada məqsəd oyun zamanı gedişlərin "fayda"sını maksimallaşdırmaq və beləliklə, oyunun nəticəsi üzərində idarəetməni əldə etməkdir.



Şəkil 10 - Tipik möhkəmləndirici maşın öyrənməsi prosesi

7.7 Transfer öyrənmə

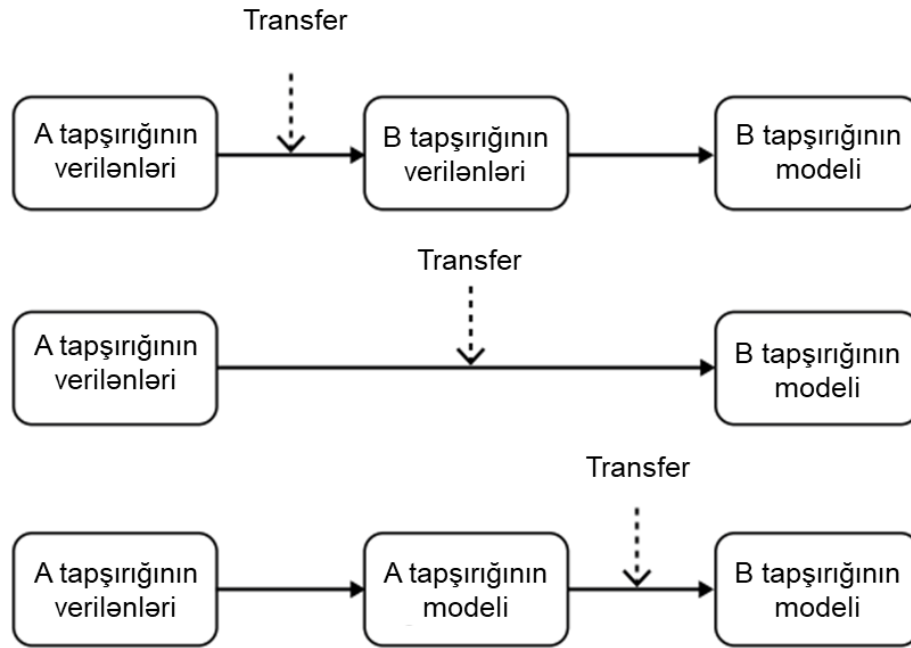
Transfer öyrənmə bir məsələni həll etmək üçün nəzərdə tutulmuş verilənlərdən əldə olunan biliklərin saxlanması və abstraktlaşdırılmasına, həmin biliklərin bu məsələ ilə zəif əlaqəsi ola bilən digər fərqli məsələnin həllində tətbiq olunmasına yönəldilmiş metodlara aiddir (məsələn, eyni məsələni tam fərqli bir tətbiq sahəsində həll etmək). Məsələn, küçə təsvirlərində ev nömrələrinin tanınması ilə əldə edilən biliklər əl ilə yazılan nömrələri tanımaq üçün də istifadə oluna bilər.

Bu öyrənmə paradigmasının tanınmış nümunəsi bir məsələni həll etmək üçün artıq öyrədilmiş modelin yeni məsələni həll etmək üçün digər məqsədlə istifadə olunduğu və əlavə təlimləndiyi incə tənzimləmədir ("fine-tuning"). Məsələn, mebel və obyektləri tanımaq üçün öyrədilmiş model yeni tətbiq sahəsindəki şəkillər üzərində təlim keçməklə müxtəlif

mühitlərdəki obyektləri (məsələn, qonaq otağı, çimərlik, mətbəx və s.) müəyyən etmək üçün incə tənzimləyə bilər.

Transfer öyrənməsini həyata keçirmək üçün nümunələr sintezi, xüsusiyyətlərə əsaslanan alqoritmlər və ya nizamlaşdırma üsulları da daxil olmaqla, digər işlənmiş üsullar da var [13].

Təcrübədə transfer öyrənmə digər öyrənmə yanaşmalarına əsaslanır və sonradan təkrar istifadə üçün biliklərin saxlanması və abstraktlaşdırılması üçün əlavə addımlarla onları tamamlayır. Şəkil 11-də göstərildiyi kimi, transfer metodları öyrənmə baş verməzdən əvvəl, öyrənmə zamanı və ya sonra tətbiq edilə bilər.



Şəkil 11 - Tipik transfer öyrənmə prosesləri

8 MAŞIN ÖYRƏNMƏSİ KONVEYERİ

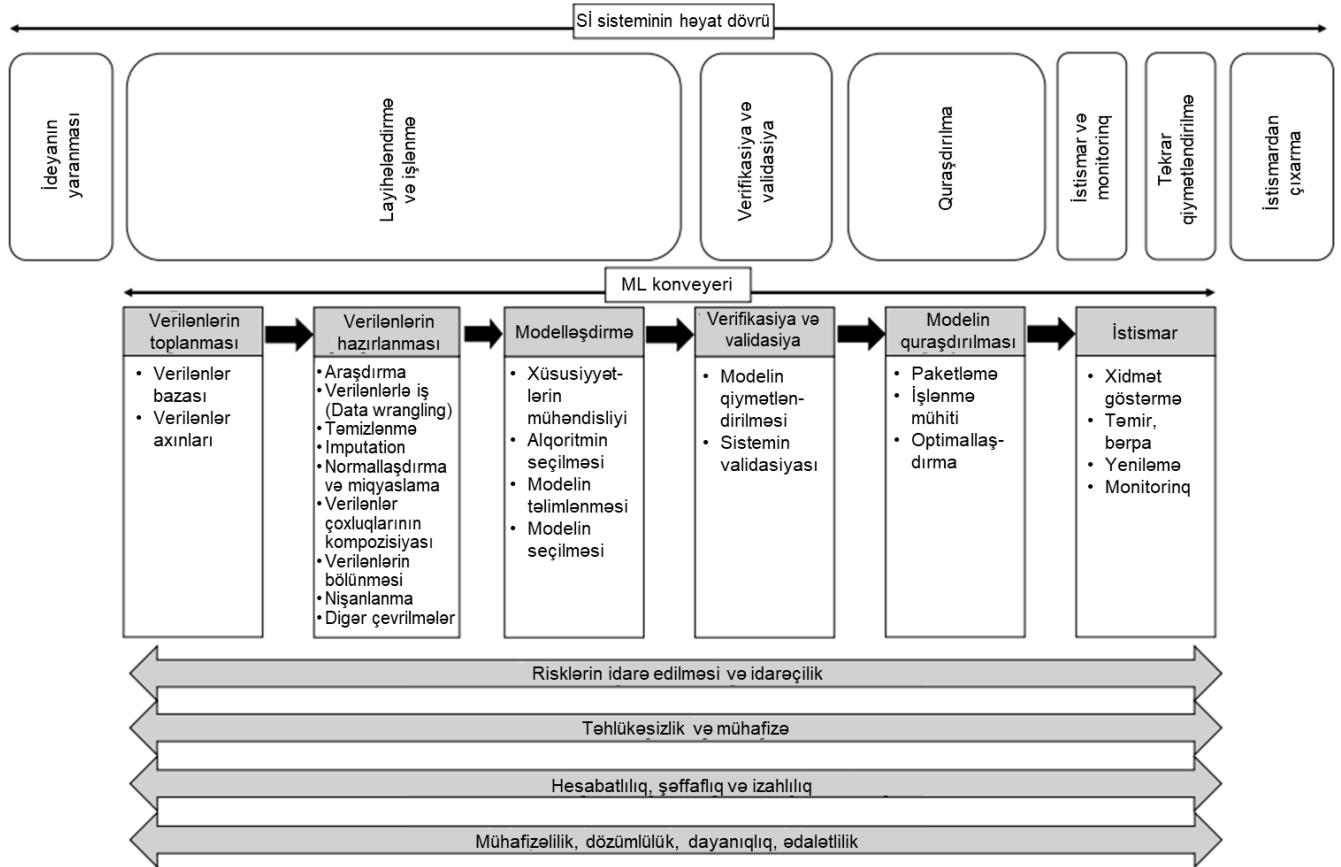
8.1 Ümumi müddəalar

Maşın öyrənməsindən istifadə edərək müəyyən məqsədə çatmaq üçün ML modeli yaradılır, qiymətləndirilir və tətbiq edilir. Bu proses adətən verilənləri, alqoritmləri və hesablama resurslarını əhatə edir. Bu bənddə hər addımda tətbiq olunan proseslər də daxil olmaqla, tipik maşın öyrənməsi konveyeri nəzərdən keçirilir. Konveyerlə işə başlamazdan əvvəl öz həllini gözləyən tapşırığı və ya məsələni müəyyən etmək lazımdır. Bu işə öz növbəsində məqsədləri və tələbləri müəyyənləşdirir. Məsələnin hərtərəfli müəyyənləşdirilməsi (məsələn, giriş və çıxış formatının dəqiq təyini də daxil olmaqla) uyğun maşın öyrənməsi alqoritmlərini seçməyə və ML modelini öyrətmək üçün lazım olan müvafiq verilənlər çoxluğunu əldə etməyə kömək edir.

Şəkil 12 ML modelinin qurulması, verifikasiyası, tətbiqi və istismarında iştirak edən maşın öyrənməsi proseslərini və onların SI sisteminin həyat dövrü mərhələləri ilə əlaqələrini

göstərir. Aşağıdakı aspektlər maşın öyrənməsi konveyeri üzrə tətbiq olunur: modelinin işlənməsi, verifikasiyası, quraşdırılması və istismarında iştirak edən konkret ML proseslərini və onların Sİ sisteminin həyat dövrü mərhələlərinə necə aid olduğunu göstərir. Bütün ML konveyeri üzrə aşağıdakı aspektlər tətbiq olunur:

- risklərin idarə edilməsi və idarəçilik;
- təhlükəsizlik və məxfilik;
- hesabatlılıq, şəffaflıq və izahlılıq;
- mühafizəlilik, dözümlülük, dayanıqlıq və ədalətlik.



Şəkil 12 — Maşın öyrənməsi konveyeri və onun Sİ sisteminin həyat dövrünə inikası

Aşağıdakı bəndlər Şəkil 12-də təqdim olunan konveyerin mərhələlərini və hər mərhələdə baş verə biləcək prosesləri təsvir edir. ML modelinin layihələndirilməsi və işlənməsi bir neçə ardıcıl və ardıcıl olmayan addımlardan ibarətdir. Proses üç addımdan ibarət olan hazırlıq mərhələsindən başlayır: verilənlər çoxluğunun toplanması və hazırlanması, xüsusiyyətlərin mühəndisliyi (əsas əlamətlərin seçilməsi) və müvafiq alqoritmin seçilməsi. Adətən, bu proseslər arasında asılılıq yoxdur və prinsip etibarlı ilə onlar istənilən ardıcılıqla və ya eyni vaxtda həyata keçirilə bilər. Bununla belə, konkret istifadə variantından asılı olaraq, bu proseslər arasında xüsusi icra qaydası və ya zəruri iterasiyalar tələb edən əlavə texniki məhdudiyyətlər ola bilər. Məsələn, verilənlər çoxluğunun təhlili daha yaxşı xüsusiyyətlərin seçilməsinə kömək edə bilər; seçilmiş xüsusiyyətlərin tərffüatları bu xüsusiyyətlərə daha çox və ya daha az ehtiyacı olan alqoritmin seçiminə təsir göstərə bilər; lakin alqoritmin seçimi

toplanmış verilənlərin annotasiyası üçün nişanların növünü və ya təfərrüatını da məhdudlaşdırıla bilər və bu halda, o ilk növbədə seçilməlidir.

Modelin faktiki yaradılması bütün hazırlıq mərhələləri tamamlandıqdan sonra baş verir. Bu, bir və ya bir neçə namizəd modelinin qurulmasını və sonra onların arasından ən yaxşı namizədin seçilməsini nəzərdə tutur.

8.2 Verilənlərin toplanması

ML modelinin qurulması, quraşdırılması və istismarında ilkin addım verilənlərin əldə edilməsidir. Verilənlər maşın öyrənməsi modelinin həyat dövrünün əsas komponentidir, çünki həm təlim, həm də modelin qiymətləndirilməsi onlardan asılıdır. Verilənlərin toplanması və hazırlanması əldə olunan verilənlərin modelin biznes məqsədlərinə cavab verməsini təmin etmək üçün xüsusi diqqət tələb edir. Təlim verilənləri çox vaxt daha böyük populyasiyanın (bu populyasiya üçün ideal representativ olan) altçoqluğudur. Müəllimlə maşın öyrənməsi modellərinin təlimlənməsi üçün istifadə edilən verilənlərə bir və ya daha çox nişan və ya məqsəd dəyişəni daxildir və onlar verilənlərin qeydi və ya nümunəsi haqqında gerçək məlumatları əks etdirir. Məsələn, təlim üçün e-poçtlar çoxluğunda hər bir e-poçt “spam” və ya “spam deyil” kimi qeyd edilə bilər.

ML modelinin dəqiqliyini və digər effektivlik göstəricilərini müəyyən etmək üçün validasiya və test verilənlərinə ehtiyac olacaq. Validasiya və test verilənləri təlim verilənləri ilə birlikdə toplanır və ya müxtəlif yollarla əldə olunan oxşar verilənlər çoxluğu ola bilər.

Lazım olan verilənlərin növü həll olunan məsələdən asılıdır. Məsələn, satışın proqnozlaşdırılması tranzaksiya verilənlərinə giriş tələb edə bilər. Şəkillərin klassifikasiyası şəkil fayllarına giriş tələb edə bilər. Eynilə, həll ediləcək məsələ verilənlər çoxluğunda hansı məlumatların (yəni hansı xüsusiyyətlərin) tələb olunduğunu müəyyənləşdirir. Məlumat mənbələri müəyyən edilir və tələb olunan verilənlər toplanır (məsələn, satın alınır və ya əldə edilir). Məlumatlar bir çox mənbədən, statik verilənlər anbarından və ya verilənlər axınından əldə edilə bilər. Təlim üçün ümumi məlumat mənbələrinə tranzaksiyalar, sensorlar, sorğular, təsvirlər, sənədlər, səslər və sosial media daxildir.

8.3 Verilənlərin hazırlanması

ML modellərinin təlimlənməsi üçün verilənlər nadir hallarda dərhal istifadəyə hazır olur. Əksər hallarda, ML modellərində onların öyrədilməsi üçün istifadə olunan verilənlərlə bağlı tələblər mövcud olur. Maşın öyrənməsi modelinin işlənməsi üçün verilənlərin hazırlanması adətən çox vaxt və səy tələb edir. Verilənlərin hazırlanması müxtəlif verilənlər çoxluğunu formatlaşdırmaq və təmizləmək üçün bir sıra avtomatlaşdırılmış alətlərdən istifadəni əhatə edə bilər. Verilənlərin hazırlanması onların konkret sistem tərəfindən istifadə edilə bilən şəkildə formatlaşdırılmasını və verilənlərin təmizlənməsini ehtiva edir. Verilənlərin təmizlənməsi onların dublikatının və yanlış verilənlərin (məsələn, həlli üçün faydasız verilənlər) silinməsini və çatışmayan verilənlərin əlavə edilməsini ehtiva edə bilər. Xətali qeydləri və ya sapmaları ehtiva edə bilən küylü verilənlər bu mərhələdə emal oluna bilər. Verilənlərin hazırlanmasına eyni zamanda verilənlərin adsızlaşdırılması (de-identifikasiya) da daxildir: əldə edilmiş verilənlər əvvəllər adsızlaşdırılsa da, bu mərhələdə verilənlərin emalı (məsələn, verilənlər çoxluğunun birləşdirilməsi) səbəbindən əlavə adsızlaşdırılma tələb oluna bilər. Bu proseslərin nümunələri verilənlərin tədqiqi, emalı, təmizlənməsi,

əvəzlənməsi (“imputation”), normallaşdırılması və miqyaslanması, verilənlər çoxluğunun tərtib edilməsi, verilənlərin bölünməsi və nişanlanması məqsədilə növbəti abzaslarda müzakirə olunur. Bu proseslər fərqli ardıcılıqla tətbiq oluna bilər.

Tədqiqat: Tədqiqat “çiy” verilənlərinin təhlilini əhatə edir. Məqsəd modelin təlimləndirilməsi üçün verilənlərin hazırlanmasında hansı proseslərin lazım olduğunu müəyyən etməyə kömək edəcək təsəvvür (məlumat) əldə etməkdir. Bu təsəvvürlərə, məsələn, verilənlər çoxluğunun forması, verilənlərin növləri, orta qiymət, ədədi verilənlərin dispersiyası və diapazonu, indeksin mövcudluğu, xüsusiyyətlər və nişanlar aiddir.

Xüsusiyyətlər üzrə verilənlərin paylanması tədqiqi verilənlərin hazırlanması ilə bağlı tədbirlərin seçilməsinə də kömək edir. Məsələn, paylanmanın təhlili verilənlərdə sapmaları, qarşılıqlı əlaqələri və nümunələri (qanunauyğunluqları) müəyyən etməyə kömək edə bilər. Bu xarakteristikalar verilənlərin hazırlanmasını təkmilləşdirilməsi üzrə müvafiq düzgün tədbirlərin seçilməsi üçün əsas kimi istifadə edilə bilər. Təlim verilənləri ilə produksiya və ya giriş verilənlərinin paylanmalarının müqayisəsi isə Sİ sisteminin maraqlı tərəflərinə planlaşdırılmış həllin effektivliyini proqnozlaşdırmağa kömək edə bilər.

Verilənlərin çeşidlənməsi: Verilənlərin yaxşı strukturlaşdırılmış və nizamlanmış verilənlər çoxluğu yaratmaq üçün yerinə yetirilir. Cədvəl verilənləri üçün verilənlərin çeşidlənməsi prosesinə sütunların birləşdirilməsi, bölünməsi və yaradılması, eləcə də verilənlər növlərinin və formatlarının dəyişdirilməsi, sətirlərdə simvolların silinməsi, dəyişdirilməsi və ya əlavə edilməsi, sütun başlıqlarının yaradılması və ya dəyişdirilməsi daxildir.

Verilənlərin təmizlənməsi: Verilənlərin təmizlənməsi verilənlərin çeşidlənməsi proseslərinə bənzərsə də, daha təfərrüatlı səviyyədə aparılır. Verilənlərin təmizlənməsi üzrə ümumi proseslərə verilənlərin dublikatının silinməsi, çatışmayan verilənlərin əlavə olunması, yanlış verilənlərin və sapmaların idarə edilməsi daxildir. Bəzi hallarda çatışmayan verilənlərin əlavə olunması və ya etibarsız verilənlərin əvəzlənməsi üçün interpolasiya həyata keçirilə bilər. Etibarsız verilənlərin nümunəsinə coğrafi “en/uzun” cütlüyü tələb olunan atributlar əvəzinə poçt indeksinin və ya ad atributu əvəzinə ədədin verilməsini göstərmək olar. Əgər verilənlərin mənasını dəyişmədən çatışmayan və ya etibarsız verilənləri əvəzləmək mümkün deyilsə, verilənlər çoxluğundan bütün yazıları və ya atributları silmək lazım gələ bilər. Bəzi maşın öyrənməsi modellərinin çatışmayan və ya etibarsız verilənlər üzərində təlimi uğursuz olur. Sapmalar verilənlərdə küy yarada bilər ki, bunun nəticəsində təlim keçmiş model quraşdırılmadan sonra produksiya verilənlərini ümumiləşdirə bilmir. Sapmalar ehtiva edən yazıları verilənlər çoxluğundan silmək zərurəti yarana bilər.

Verilənlərin əvəzlənməsi: Verilənlərin əvəzlənməsi çatışmayan verilənlərin əvəzinə hesablanmış qiymətlərdən istifadə edən təmizləmə prosesidir. Əvəzlənmə üsullarına sadə və ya çoxsaylı əvəzlənmə daxildir. Sadə əvəzlənmə (məsələn, “hot-deck”, “cold-deck”, reqressiya və orta qiymət) çatışmayan qiymətlərlə bağlı əhəmiyyətli qeyri-müəyyənliyin olmaması şərti ilə istifadə üçün birbaşa hədəf qiymətini generasiya edir. Çoxsaylı əvəzlənmə isə əvəzlənmə modelində çatışmayan qiymətlərlə bağlı qeyri-müəyyənliyin mövcudluğunu əks etdirir və bu zaman bir az fərqli modellərdən təsadüfi olaraq seçilmiş bəzi əvəzlənmiş qiymətləri nəzərə almaqla, hədəf qiymətini generasiya edir.

Normallaşdırma və miqyaslama: Maşın öyrənməsi modeli təlimləndirərkən, verilənlər çoxluğunda ədədi xüsusiyyətlər diapazonundakı əhəmiyyətli fərqlər xüsusiyyətlərin müqayisəsini mümkün etməz edə bilər. Normallaşdırma, bir tərəfdən, ayrıca verilənlər

nümunələrini vahid normaya (0 və 1 aralığında) miqyaslamaq (inikas) üçün istifadə edilə bilər. Digər tərəfdən, miqyaslama ayrıca nümunələr əsasında normal paylanma yaratmaq və ya nümunələri müəyyən bir aralığa sıfırdırmaq üçün istifadə edilə bilər.

Modelin nəticələrinin uyğunluğunu təmin etmək üçün verilənlərə miqyaslama və ya normallaşdırma tətbiq oluna bilər.

Verilənlərin normallaşdırılması üsullarına nümunələr aşağıdakılardır:

- Min-max normallaşdırma: verilənlərin minimum və maksimum qiymətlərə (çox vaxt 0 və 1) uyğun xətti çevrilməsi;
- Z-normallaşdırması: verilənlər orta və standart sapmalara əsasən miqyaslanır;
- Onluq miqyaslama: atributların qiymətlərinin onluq nöqtəsinin yerinin dəyişdirilməsi.

Verilənlər çoxluğunun tərtib olunması: Verilənlər çoxluğunun tərtib edilməsi müxtəlif verilənlər mənbələrindən toplanmış verilənləri vahid verilənlər çoxluğunda cəmləşdirərkən onların seçilməsi və ya kompilyasiyasını ehtiva edir. Adətən bu kompilyasiya təlim keçəcək xüsusiyyətlər, habelə nəzərdə tutulan xüsusiyyətlərin öyrədilməsi üçün obyektlərin zəruri paylanması və təqdimatı nəzərə alınmaqla aparılır.

Verilənlər çoxluğunun bölünməsi: ML modelləri üçün təlim verilənləri, həmçinin modelin seçimi üçün validasiya verilənləri və modelin qiymətləndirilməsi üçün test verilənləri tələb olunur. Təlim, validasiya və test verilənləri kəsişməməlidir. Bəzi hallarda onlar ayrıca şəkildə toplanır. Lakin elə olur ki, onlar eyni mənbə(lər)dən birgə alınır, bu halda mövcud verilənlər ayrı-ayrı çoxluqlara bölünməlidir. Adətən verilənlərin validasiya və test hissələri təlim hissələrindən çox kiçik seçilir.

Nişanlama: Verilənlərin nişanlanması nümunəyə məqsəd dəyişəninin təyin edilməsi (mənimsədilməsi) prosesidir. Verilənlərin müvafiq nişanlarla əlaqələndirilməsi etalon verilənləri müəyyən etməyə imkan verir. Verilənlərin nişanlanması nümunədən məna kəsb edən faydalı məlumatların (məsələn, təsvirdəki insanların sayı) əldə edilməsinə əsaslanıla bilər. Həmçinin nişan yaratmaq üçün nümunə öz məzmunu haqqında məlumatla (məsələn, şəkildə təsvir edilən iki nəfərin mübahisə edib-etməməsi) zənginləşdirilə bilər. Verilənlər insanlar tərəfindən əl ilə və ya bəzi hallarda kompüter tərəfindən avtomatik olaraq nişanlanıla bilər.

Nümunə verilənlərinin nişanlanması müəllimlə öyrənmədə olduğu kimi təlim üçün lazım ola bilər. Nişanlama hətta müəllimlə öyrənmə olmadıqda da yaradılan modellərin qiymətləndirməsi üçün tələb oluna bilər. Belə hallarda, müvafiq verilənlər onların qeydi və ya nümunəsi haqqında etalon verilənləri əks etdirən bir və ya bir neçə nişanı və ya məqsəd dəyişənini ehtiva edir.

Adsızlaşdırma: ISO/IEC 20889 standartında təsvir olunduğu kimi, adsızlaşdırma verilənlərin subyektii ilə identikləşdirən atributlar çoxluğu arasında əlaqənin silinməsi prosesidir. Əgər verilənlər çoxluğunda fərdi identikləşdirən məlumatlar (PII) varsa, o zaman PII-ni silmək tələb oluna bilər. Xüsusilə, ML modelinə təlim keçmək məqsədilə tibbi məlumatlardan istifadə etmək üçün fərdi məlumatların və ya sağlamlığa dair mühafizə olunan məlumatların (PHI) adsızlaşdırılması tələb edilə bilər. Əlavə olaraq, təlim verilənlərinin zərərli və ya arzuolunmaz təsirlərə məruz qala biləcək verilənlərlə

çirklənməsinin qarşısını almaq üçün verilənlərin “zəhərlənməsi”nin (“poisoning”) yoxlanılması çox vacibdir. Təlim verilənləri çoxluğu tam və kifayət deyilsə və ya təlim üçün tələb olunan paylanmanı təhrif edirsə, səhvən təqdim olunarsa, proqnozlaşdırılan nəticələr qərəzli (qeyri-obyektiv) və ya diskriminativ ola bilər.

8.4 Modelləşdirmə

Modelləşdirmə quraşdırılmağa hazır olan, təlim və testdən keçmiş ML modelinin işlənməsi prosesidir.

Xüsusiyyətlər mühəndisliyi (əsas əlamətlərin seçilməsi): Xüsusiyyətlər (adətən, cədvəl verilənlərində sütun başlıqları kimi təqdim olunur) verilənlər çoxluğundakı elementlərin kəmiyyət deskriptorlarıdır. Xüsusiyyətlər ədədi, mətn tipli, kəsilməz və ya kateqorial ola bilər. Kateqorial xüsusiyyətlər nominal və ya ordinal (sıralanan) ola bilər. Tarix, vaxt, ad, ünvan, yaş və satış məbləği xüsusiyyətlərə nümunədir. Xüsusiyyətlər mühəndisliyi ML modelində istifadə üçün xüsusiyyətlərin seçilməsi, xarakterizə edilməsi və optimallaşdırılması prosesidir. Xüsusiyyətlər mühəndisliyi aşağıdakıları ehtiva edir:

- **Kodlaşdırma:** Effektivlik baxımından və rahat istifadə üçün mətn tipli və kateqorial xüsusiyyətlər əksər hallarda ədədi identifikatorlara çevrilir;
- **Verilənlər növünün çevrilməsi:** Verilmiş modelin tələblərinə cavab vermək üçün xüsusiyyətin verilənlər növünün çevrilməsi lazım ola bilər;
- **Ölçülərin azaldılması:** Giriş verilənlərinin ölçüsünün böyük olması bəzi ML alqoritmləri üçün çətinlik yarada bilər. Buna görə də hər nümunənin atributlarının sayını azaltmaq lazım gələ bilər (çox vaxt növbə ilə başqa ML modelindən istifadə etməklə). Həmçinin bax 6.2.6;
- **Xüsusiyyətlərin seçilməsi:** Xüsusiyyət seçiminin təməlinə verilənlərin lazımsız (izafi) və ya uyğun olmayan xüsusiyyətləri ehtiva edə bilməsi və buna görə də onların silinə bilməsi durur. Müəyyən xüsusiyyətlər modeldə küy yaradır və ya modelin proqnozlarını təhrif edir, buna görə də onlar silinməlidir. Düzgün xüsusiyyətləri seçmək ML modelinin məhsuldarlığı üçün çox vacib ola bilər. Aşağıdakı səbəblər də nəzərə alınmaqla, (lakin bunlarla məhdudlaşmayaraq) mövcud verilənlər çoxluğundan dəyişənlərin atributlarını seçmək lazımdır:
 - interpretasiyasını asanlaşdırmaq üçün modellərin sadələşdirilməsi;
 - təlim vaxtının daha da qısaldılması;
 - uyğun ölçünün təmin edilməsi. Müəyyən xüsusiyyətlər modeldə küy yaradaraq və ya modelin generasiya etdiyi proqnozları təhrif edərək, çox sayda xüsusiyyətlərə malik verilənlər “ölçü lənəti” (“curse of dimensionality”) ilə qarşılaşa bilər;
 - modelin həddən artıq təlimlənməsinin azaldılması.

Faydalı xüsusiyyətləri optimal şəkildə seçmək üçün istifadə edilə bilən bir neçə xüsusiyyət seçimi alqoritmı var. Sadə yanaşma problemin həllinə intuitiv olaraq uyğun olmayan xüsusiyyətləri istisna etməkdir. İkinci üsul modelin performansına hansı xüsusiyyətlərin az təsir göstərdiyini görmək üçün verilənlərin vizuallaşdırılmasıdır. Bəzi hallarda, performans fərqi görmək üçün modeli konkret xüsusiyyətlərlə və onlarsız təlimləndirmək və test etmək lazım gələ bilər. Qərarlar ağacı kimi bəzi ML alqoritmləri öyrənmə prosesinin bir hissəsi kimi xüsusiyyət seçimini ehtiva edir.

Alqoritm seçimi: Konveyerin bu mərhələsində tapşırığa uyğun ML alqoritmni seçmək lazımdır (məsələn, klassifikasiya, reqressiya, klasterizasiya, tövsiyə). ML alqoritmləri ML modelinin yaradılmasının əsasını təşkil edir. Hər növ ML tapşırığı üçün bir neçə alqoritm mövcuddur və daima yeni üsullar da meydana çıxır. Kitabxanalar və ya ML alqoritmlərinin kolleksiyaları asanlıqla əldə edilə bilər. Bəzi hallarda, yeni alqoritm yaratmaq və ya mövcud alqoritmni modifikasiya etmək ya zəruri, ya da arzuolunan olur. Müəyyən məsələni və ya tapşırığı həll etmək üçün kitabxanadan bir və ya bir neçə alqoritm seçilə, yaxud alqoritm yaradıla bilər. Kitabxanadakı alqoritmlərin tətbiqi qaydaları kitabxana sənədlərində təsvir olunur. Ümumi halda elmi və müqayisə üçün ədəbiyyat istifadə üçün əlçatan olur.

Həmçinin son proqnozun hər bir modelin proqnozuna tətbiq edilən funksiyaya əsaslandığı modellər ansamblından da istifadə etmək mümkündür.

Model təlimi: Modeldə hər bir parametrlər üçün sabit göstəricilərin və ya çəkirlərin müəyyən edilməsi üçün ML modeli təlim verilənləri üzərində iterativ olaraq təlim keçir. Təlim modelin “uyğunlaşdırılması” da adlanır. Təlimin məqsədi istehsal verilənləri üzərində yaxşı ümumiləşdirmə imkanlarına malik model yaratmaqdır. Model təlim verilənlərini yadda saxlayır, lakin yaxşı ümumiləşdirmə aparmırsa, həddən artıq təlimlənmə baş verir; bu, kifayət qədər təlim nümunələri olmadıqda baş verə bilər. Xüsusiyyətlər düzgün seçilmədikdə və ya çoxsaylı təlim nümunələri üzərində düzgün öyrənmək üçün kifayət qədər model parametrləri olmadıqda, yetərsiz öyrənmə baş verə bilər. Təlim prosesi, ilk növbədə, layihələndirmə və işlənmə mərhələsində tətbiq olunur, lakin, adətən, modelin istifadəsi zamanı təkrar təlim və ya fasiləsiz təlim üçün daha sonra da yenidən həyata keçirilə bilər. Təcrübədə modelin istifadəsi zamanı ona fasiləsiz təlim keçmək olar; lakin bu proses həm əvvəlki istifadə olunan, həm də yeni verilənlər üzərində davamlı yaxşı məhsuldarlığın təmin edilməsində böyük çətinliklərə səbəb olur.

Model seçimi: Hiperparametrlərin sazlanması kimi də tanınan model seçimi ən yaxşı modeli seçmək üçün ən yaxşı nəticələri təmin edən qiymətləri müəyyən etmək məqsədilə hiperparametrləri qiymətləndirmək və optimallaşdırmaq üçün validasiya verilənləri çoxluğundan istifadə prosesidir. Hiperparametr sazlanmasını avtomatlaşdırmaq üçün şəbəkə (“grid”) üzrə müfəssəl axtarış və təsadüfi parametrlər optimallaşdırması kimi üsullardan istifadə etmək olar. Bundan əlavə, bəzi metodlar verilmiş parametrlər üzrə qiymətlər şəbəkəsindən istifadə edərək modeli təlimləndirmək və qiymətləndirmək imkanına malikdir.

Verilənlər azdırsa, həmçinin təlim və validasiya üçün ayrıca verilənlər isə lap azdırsa, çarpaz qiymətləndirmədən istifadə olunur. Hiperparametrin sazlanması üçün tez-tez “k-fold” çarpaz validasiyasından istifadə olunur. “K-fold” çarpaz validasiyasında zamanı ilkin təlim verilənləri çoxluğu k altçoxluğuna (altqatlarına, “folds”) bölünür. Hər bir qat üçün model digər qatların birləşməsində təlim keçir və sonra onun çıxış verilənlərinin xətası bu qatdan istifadə etməklə qiymətləndirilir. Hər bir qatın yalnız bir nümunədən ibarət olduğu xüsusi hal “birini istisna et” (“leave-one-out”) çarpaz-validasiyası adlanır. Çarpa-validasiya məlumat sızması səbəbindən həddən artıq öyrənmənin qarşısını almağa kömək edə bilər.

Mallow's Cp metodu, Akaike və Bayes informasiya meyarları kimi üsullar da aşkar şəkildə hesablama aparmadan fərqli modellərin performansını müqayisə etmək üçün istifadə edilə bilər.

8.5 Verifikasiya və validasiya

Modelin qiymətləndirilməsi: Modelin qiymətləndirilməsi test verilənləri üzrə modelin generasiya etdiyi proqnozların verilənlərdəki faktiki nişanlarla müqayisəsi prosesidir. Təlim keçmiş modellərin performansını qiymətləndirmək üçün bir çox metrikalar mövcuddur. Bu göstəricilər 6.5.5-ci yarımbənddə təsvir olunur.

İşlənmə və istifadə zamanı müəyyən edilmiş risklərin azaldılması şərtilə ilkin layihələndirmə məqsədlərinə ən çox uyğun gələn modeli təyin etmək üçün adətən bir neçə modelin test edilməsi adi haldır. Lakin belə təcrübə xüsusi ehtiyatla tətbiq edilməlidir ki, test verilənləri üzərində yerinə yetirilməməli olan model seçimi prosesindən fərqlənsin. Müqayisəli qiymətləndirmə yalnız bir-birindən çox fərqlənən (məsələn, təkcə fərqli hiperparametrlər deyil, fərqli təlim verilənləri, fərqli ML alqoritmləri və ya fərqli ML yanaşmalarına əsaslanan) bir neçə model üçün aparılmalıdır. Əks halda, hiperparametrlərin test verilənlərinə həddən artıq uyğunlaşdırılması riski yaranır, bu da faktiki olaraq modelin performansının olduğundan yüksək qiymətləndirilməsi deməkdir.

Sistemin validasiyası: Sistemin validasiyası - təlim keçmiş və quraşdırılmış modelin maraqlı tərəflərin bütün ehtiyaclarına cavab verdiyini və xüsusiyyətlərin əldə edildiyi mühit haqqında ilkin ehtimalların düzgün olduğunu yoxlamaqdır. Bu proses layihələndirmə və işlənmə mərhələsində hiperparametrlərin sazlanması üçün istifadə olunan validasiya verilənləri ilə əlaqəli deyil.

8.6 Modelin quraşdırılması

ML modeli təlim keçdikdən və qiymətləndirildikdən sonra produksiya verilənləri əsasında proqnozlar vermək üçün istifadə edilə biləcək istismar mühitində quraşdırılmalıdır. ML modelləri veb-tətbiqlər (REST API vasitəsilə əlçatanlıq təmin olunur) və ya yerli ("native") tətbiqlər kimi quraşdırıla bilər.

Paketləşdirmə: Təlim keçmiş ML modelləri ənənəvi olaraq serverlərdə, klasterlərdə və ya virtual maşınlar quraşdırılır. ML modellərini təqdim etmək və ya mübadilə etmək üçün açıq neyron şəbəkələri mübadiləsi kitabxanası (ONNX) [15] və neyron şəbəkələri mübadiləsi formatı (NNEF) [16] kimi bir neçə açıq format mövcuddur. Son vaxtlar quraşdırma prosesi əhəmiyyətli dərəcədə sadələşdirilərək, modellər bütün asılılıq amilləri nəzərə alınaraq paketlənmiş, işə hazır konteynerlərdə quraşdırılır. Qurğuların emal gücünün və yaddaş həcminin sürətlə artması təlim keçmiş ML modellərinin pilotsuz uçuş aparatları və dronlar kimi cihazlarda da quraşdırılmasını mümkün edir.

İcra mühiti: R və ya Python kimi yüksək səviyyəli dillər ML modellərinin işlənməsində tez-tez istifadə olunur. Daha sonra təlim keçmiş model (məsələn, interpretator və ya "just-in-time" kompilyator kimi) müvafiq abstraksiya səviyyəsinə malik olan istənilən icra mühitində quraşdırıla bilər. Modelin icra mühiti onun işlənmə mühitindən daha aşağı səviyyəli ola bilər. İcra mühitinə və onun işlənmə mühitindən fərqlənə biləcəyinə xüsusi diqqət yetirilməlidir.

Optimallaşdırma: Verilənlərin növlərini dəqiqdən təqribiyə və ya maşından asılı olmayanlardan hədəf platformada dəstəklənən növlərə (məsələn, neyron şəbəkəsinin çəkirlərinin tam qiymətlərlə ifadə olunması) qədər diapazonda dəyişdirməklə model hədəf platforması üçün optimallaşdırıla bilər. Optimallaşdırma səmərəli aparat realizasiyası mövcud olan yerdəyişmə və ya toplama əməliyyatları vasitəsilə də həyata keçirilə bilər. Optimallaşdırmanın növündən və dərəcəsindən asılı olaraq optimallaşdırılmış modelin yeni

qiymətləndirilməsi zərurəti yarana və buna görə də verifikasiya və validasiya mərhələlərinə qayıtmaq tələb oluna bilər.

8.7 İstismar mərhələsi

Təlim modeli quraşdırıldıqdan sonra müstəqil tətbiqi proqram və ya digər tətbiqlərə inteqrasiya funksiyası kimi istifadə edilə bilər. ML modelləri sabit produksiya verilənləri çoxluğu ilə paket rejimdə və ya real vaxt produksiya verilənləri axınları ilə fasiləsiz rejimdə istifadə olunur. Modelin çıxış verilənləri tətbiqə geri göndərilə və ya bəzi hallarda insan tərəfindən interpretasiya edilə bilər.

Texniki xidmət və bərpa: İstismar mərhələsində ML modeli və müvafiq tətbiq üçün digər Sİ sistemləri və ya adi proqram sistemlərinə analogi olaraq, texniki xidmət və bərpa prosedurları tələb olunur. Bu mərhələdə ML-in spesifikasiyası yeniləmə və monitoring prosesləri ilə əlaqədardır.

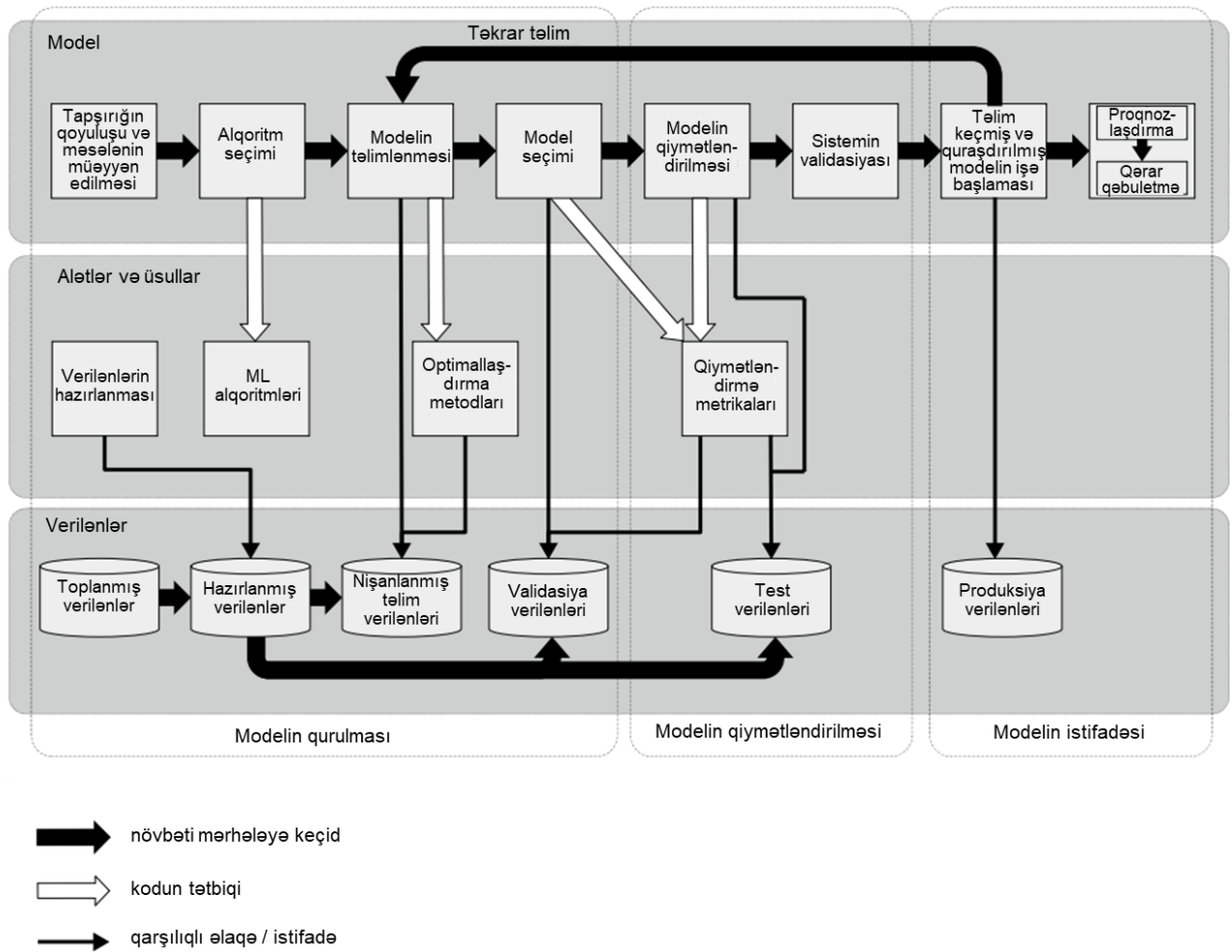
Yeniləmə: Təkrar təlim prosesi cari produksiya verilənlərinin xüsusiyyətlərini əks etdirən təlim verilənlərindən istifadə etməklə həyata keçirilir. Yeniləmələr həmçinin alqoritmik dəyişiklikləri də əhatə edə bilər ki, bu da (dəyişiklikdən asılı olaraq) yalnız qismən təkrar təlim və ya yeni modeli sıfırdan öyrətmək tələb edir.

Monitoring: ML modelinin icrası adətən istismar mərhələsində izlənilir, çünki verilənlərin dreyfi, konsepsiya dreyfi və ya zəif ümumiləşdirmə kimi amillər performansı təsir edə bilər. Fasiləsiz öyrənmə zamanı əlavə dəstək tədbirlərinin göstərilməsinə ehtiyac yarana və monitoring ML modelinin yenidən əlavə qiymətləndirmələrini əhatə edə bilər.

8.8 ML konveyeri əsasında maşın öyrənməsi prosesinə dair nümunə

ML konveyer prosesləri fərqli öyrənmə üsullarına və ML modellərinin işlənməsinə tətbiq edilə bilər. Bu bənddə proseslər müəllimlə maşın öyrənməsi üçün tətbiq edilir.

Şəkil 13 ML sisteminin müxtəlif altkomponentlərini və bu proseslərin bir-biri ilə necə uyğunlaşa biləcəyini göstərir. Nəzərə alınmalıdır ki, ML konveyerində təsvir olunan proseslərin heç də hamısı şəkil 13-də (sadələşdirilmiş nümunə olduğu üçün) təsvir edilməyib. Bəzi proseslər fərqli ardıcılıqla yerinə yetirilə və ya paralel aparıla bilər. Bundan əlavə, modelin işlənməsinin müxtəlif mərhələlərində alətlərdən istifadə, eləcə də onların verilənlərdən asılılığı təsvir edilir. Modelin işlənməsi üçün alətlərin istifadəsi ilə verilənlər üzərində proseslərin icrası arasında fərq diqqətə çatdırılır.



Şəkil 13 — ML konveyeri əsasında müəllimlə maşın öyrənməsi prosesi

Şəkil 13-də modelin təsvirinin yuxarı hissəsində müəllimlə öyrənmədən istifadə edən sistemin proseslərinin addımları təsvir edilir. Orta hissədə (Alətlər və üsullar) prosesin addımlarının yerinə yetirilməsi üçün tələb olunan alətlər və verilənlərin emalı elementləri, aşağı hissədə isə verilənlərin bütün konveyer üçün tələb olunan müxtəlif elementləri yerləşir.

“Modelin qurulması” ilkin olaraq məsələnin həlli üçün məqsəd və tələbləri müəyyən edən “məsələnin müəyyən edilməsi” və “tapşırığın qoyuluşu” addımlarını əhatə edir. Daha sonra tələb olunan həlli təmin etmək üçün ən uyğun ML metodunu seçməyi hədəfləyən “alqoritm seçimi” addımına keçid edilir. “Modelin təlimlənməsi” addımı nişanlanmış təlim verilənlərindən istifadə etməklə təlim keçmiş namizəd-modelin qurulmasını nəzərdə tutur. Namizəd-modellər arasında, nəticə etibarilə, yekun ML modelini seçmək üçün validasiya verilənləri üzrə qiymətləndirmə metrikaları hesablanır.

“Modelin qiymətləndirilməsi” mərhələsi təlim keçmiş modelin performansını qiymətləndirmək üçün test verilənlərinə qiymətləndirmə metrikalarını tətbiq edir. Bu addımlar kifayət etdikdə, “sistemin validasiyası” aparılır.

“Modelin istifadəsi” təlim keçmiş modelin quraşdırılmasını təmin edən “təlim keçmiş və quraşdırılmış modelin işə başlaması” addımını ehtiva edir, belə ki, model ML tətbiqi proqramı üçün produksiya verilənləri əsasında işləyir. Modelin produksiya verilənləri əsasında işləməsi sistemin proqnozlar verməsi və modelin çıxış verilənlərinə əsasən qərarlar qəbul etməsi ilə nəticələnir.

Modelin təkrar təlimi fasiləsiz öyrənmə vasitəsilə və ya istifadədə olan təlim keçmiş modelin performansının qiymətləndirilməsinə əsasən aparıla bilər. Performans sistem üçün müəyyən edilmiş tələblərə cavab vermədikdə, performansın qiymətləndirilməsi bəzi produksiya verilənləri üçün əlavə təlimə ehtiyac olduğunu göstərə bilər.

Şəkil 13 əlavə olaraq, “Əlavə A”da təqdim olunduğu kimi, verilənlər axınlarını və onlarla əlaqəli verilənlərdən istifadə üzrə təlimatları təsvir etmək üçün istifadə edilə bilər. Bu cür təlimatlar ISO/IEC 19944-1 standartında təsvir edilən taksonomiya və verilənlərdən istifadə üzrə təlimatları formatını ehtiva edir.

ƏLAVƏ A (informativ)

Müəllimlə öyrənmə prosesi üçün verilənlər axını nümunəsi və verilənlərdən istifadə üzrə təlimatlar

A.1 Ümumi

Bu əlavədə verilənlər axınlarını təsvir etmək və ML üçün verilənlərdən istifadə üzrə təlimatlar hazırlamaq məqsədilə ISO/IEC 19944-1 standartının istifadəsi nümayiş etdirilir. Bu əlavənin məzmunu nümunə üçündür və lazım gəldikdə, ona əlavələr edilə bilər. ISO/IEC 19944-1 standartını ML proseslərinə tətbiq etməklə Sİ sistemlərində hesabatlılığın və şəffaflığın təmini üçün zəruri olan struktur və fundamental elementlər işlənə bilər.

A.2 Müəllimlə maşın öyrənməsi prosesində verilənlər axını

A.2.1 Ümumi müddəalar

Burada Şəkil 13-də göstərilən ML prosesi əsasında verilənlər axınının nümunəsi təsvir olunur.

Qeyd edək ki, A.2.2-də verilənlər axını “bulud” hesablamalarının etalon arxitekturası (ISO/IEC 17789) çərçivəsində funksional komponentlər arasındakı axınlar kimi təsvir edilmir, əvəzində ML konveyerinin və ML modelinin işlənməsi proseslərinin müxtəlif mərhələlərindəki verilənlər axınları kimi təqdim olunur.

A.2.2 Verilənlər axını təsvirləri

A.2.2.1 Verilənlər axını 1

Aşağıdakı verilənlər axını şəkil 13-də identifikasiya oluna və aşağıdakı kimi təsvir edilə bilər:

Toplanmış verilənlər → hazırlanmış verilənlər

ML üçün verilənlər çoxluğunu yaratmaq məqsədilə verilənlər bir və ya daha çox mənbədən toplanır. Verilənlər çoxluğunu ML konveyerinin model təlimi mərhələsi üçün hazır vəziyyətə gətirmək məqsədilə bu çoxluğa çoxsaylı proseslər tətbiq edilir.

Bu proseslər 8.3-cü bənddə təsvir olunur.

A.2.2.2 Verilənlər axını 2

Aşağıdakı verilənlər axını şəkil 13-də identifikasiya oluna və aşağıdakı kimi təsvir edilə bilər:

Hazırlanmış verilənlər → nişanlanmış təlim verilənləri

Müəllimlə maşın öyrənməsi üçün məqsəd dəyişəninin etalon qiyməti göstərilməlidir. Məqsəd dəyişəninin qiyməti onun nişanı kimi tanınır. Təcrübədə bir proksidən (vasitəçi) etalon qiyməti kimi istifadə etmək olar, digər proksilər səhv və qərəzli nəticə təqdim edə bilər.

A.2.2.3 Verilənlər axını 3

Aşağıdakı verilənlər axını Şəkil 13-də identifikasiya oluna və aşağıdakı kimi təsvir edilə bilər:

Təlim keçmiş və quraşdırılmış modelin işə salınması → proqnozlaşdırma

Model quraşdırıldıqdan və işə salındıqdan sonra produksiya verilənləri üzərində tətbiq edilir. Model (adətən onun çıxışlarına təyin edilmiş ehtimallarla) çıxış verilənləri generasiya edir. Bu çıxış verilənləri öz növbəsində ya avtomatlaşdırılmış qərar qəbul etmədə, ya da insanlar tərəfindən qərar qəbul edilərkən tövsiyələr kimi nəzərə alınması üçün istifadə edilə bilər.

A.3 Maşın öyrənməsində verilənlərdən istifadə

A.3.1 Ümumi müddəalar

Bu bənddə ISO/IEC 19944-1:2020 standartının 10.2-ci bəndində təsvir edilən verilənlərdən istifadə üzrə təlimatlar strukturu əsasında verilənlərdən istifadə üzrə nümunələr yaradılır. Aşağıda A.2.2-də təsvir edilmiş və Şəkil 13-də göstərilən verilənlər axınlarına əsaslanan müəllimlə maşın öyrənməsi çərçivəsində verilənlərdən istifadə üzrə nümunələr əks olunur.

A.3.2 Verilənlərdən istifadə üzrə nümunə təlimat A

Bu nümunə sığorta şirkəti tərəfindən maşın öyrənməsindən istifadəni nəzərdən keçirir. Bənd A.2.2.1-də (verilənlər axını 1) təsvir edilən verilənlər axını çərçivəsində verilənlərin istifadəsi üzrə aşağıdakı təlimat nümunəsini istifadə etmək olar:

NÜMUNƏ – Verilənlərin emalı alətləri ML modelinin işlənməsi üçün toplanmış verilənləri təkmilləşdirmək və hazırlamaq məqsədilə ev sığortası ilə bağlı ərizələrdəki anonimləşdirilmiş maliyyə məlumatlarından istifadə edir.

A.3.3 Verilənlərdən istifadə üzrə nümunə təlimat B

Bu nümunə avtomatlaşdırılmış yol nəqliyyatı obyektlərinin tanınması sistemlərində maşın öyrənməsinin istifadəsini nəzərdən keçirir. Bənd A.2.2.2-də (verilənlər axını 2) təsvir edilən verilənlər axını çərçivəsində verilənlərin istifadəsi üzrə aşağıdakı təlimat nümunəsini istifadə etmək olar:

NÜMUNƏ – “Model öyrənmə” prosesi obyektləri tanımaq məqsədilə ML modelini öyrətmək üçün avtomobilin bort kameralarından əldə edilmiş və nişanlanmış sensor məlumatlarını istifadə edir.

A.3.4 Verilənlərdən istifadə üzrə nümunə təlimat C

Bu nümunə ML-in təyyarə mühərriki üçün proqnozlaşdırılan texniki xidmət sistemlərində istifadəsini nəzərdən keçirir. Belə sistemlərdə mühərrikin işinə nəzarət üçün kömək edən temperatur sensorları var. Bənd A.2.2.3-də (verilənlər axını 3) təsvir edilən verilənlər axını çərçivəsində verilənlərin istifadəsi üzrə aşağıdakı təlimat nümunəsini istifadə etmək olar:

NÜMUNƏ – Mühərrikə quraşdırılmış temperatur sensorlarından daxil olan verilənlər ML modeli tərəfindən komponentlərin potensial nasazlıqlarını proqnozlaşdırmaq və mühərrikin işini (xarakteristikalarını) və texniki xidmətini yaxşılaşdırmaq üçün istifadə olunur.

ƏDƏBİYYAT

- [1] Stuhlsatz A., Lippel J., Zielke T., Feature Extraction with Deep Neural Networks by a Generalized Discriminant Analysis. *IEEE Transactions on Neural Networks and Learning Systems*. 2012, **23** (4), 596—608. doi: 10.1109/tnnls.2012.2183645
- [2] MacLennan B.J., Connectionist Approaches. *International Encyclopaedia of the Social & Behavioral Sciences*. 2001, 2568-2573. doi: 10.1016/b0-08-043076-7/00537-4
- [3] Rosenblatt F., The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*. 1958, Nov, **65**(6):386-408. doi: 10.1037/h0042519
- [4] Rumelhart D., Hinton G., Williams R., Learning representations by back-propagating errors. *Nature* 1986, **323**, 533–536. doi: 10.1038/323533a0
- [5] Elman Jeffrey L., Finding structure in time. *Cognitive science*. 1990, **14.2**, 179-211. doi: 0.1016/0364-0213(90)90002-E
- [6] Hochreiter Sepp, Schmidhuber Jürgen, Long short-term memory. *Neural computation*. 1997, **9.8**, 1735-1780. doi: 10.1162/neco.1997.9.8.1735
- [7] Lecun Y, Bottou L, Bengio Y, Haffner P, Gradient-based learning applied to document recognition. *Proceedings of the IEEE*. 1998, **86**(11), 2278-2324. doi: 10.1109/5.726791
- [8] Heckerman D., A Tutorial on Learning with Bayesian Networks. *Innovations in Bayesian Networks. Studies in Computational Intelligence*, **156**. Berlin, Heidelberg: Springer. 2008. doi: 10.1007/978-3-540-85066-3_3
- [9] Fawcett T., ROC graphs: Notes and practical considerations for researchers. *Machine learning*. 2004, **31**, 1—38.
- [10] Delgado R, Tibau X. A., Why Cohen's Kappa should be avoided as performance measure in classification. *PloS one*. 2019, **14**(9). doi: 10.1371/journal.pone.0222916.
- [11] Giles M., Foody. Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote Sensing of Environment*. 2020, **239**. doi: 10.1016/j.rse.2019.111630
- [12] ISO/IEC 19944-1:2020, *Cloud computing and distributed platforms — Data flow, data categories and data use — Part 1: Fundamentals*
- [13] Yang Q., Zhang Y., Dai W., Pan S., *Transfer Learning*. Cambridge: Cambridge University Press. 2020. doi: 10.1017/9781139061773
- [14] ISO/IEC 20889, *Privacy enhancing data de-identification terminology and classification of techniques*
- [15] *Open neural network exchange*. <https://onnx.ai/>
- [16] *Neural network exchange format*. <https://www.khronos.org/nnef>
- [17] ISO/IEC 17789, *Information technology — Cloud computing — Reference architecture*