

AZƏRBAYCAN RESPUBLİKASI TƏHSİL NAZİRLİYİ

AZƏRBAYCAN DÖVLƏT İQTİSAD UNİVERSİTETİ

«MAGİSTRATURA MƏRKƏZİ»

Əlyazması hüququnda

Xələfli Sara Bəhruz qızı

**MÖVZU: «İQTİSADİ GÖSTƏRİCİLƏRİN PROQNOZLAŞDIRILMASI
ÜÇÜN MƏLUMATLARIN İNTELLEKTUAL TƏHLİLİ MODELİNİN
QURULMASI »**

MAGİSTR DİSSERTASIYASI

İxtisasın şifri və adı

060509

“Kompüter elmləri”

İxtisaslaşmanın adı

“İqtisadi informasiya sistemləri”

Elmi rəhbər:

t.e.n. dos. R.H.Gülməmmədov

Magistr proqramının rəhbəri:

akad. Ə. M. ABBASOV

KAFEDRA MÜDİRİ:

akad. Ə. M. ABBASOV

BAKİ 2020

MÜNDƏRİCAT

GİRİŞ	3
FƏSİL I. İQTİSADI GÖSTƏRİCİLƏRİN MONİTORİNGİNİN METOD VƏ SİSTEMLƏRİ.....	7
1.1. İqtisadi sistemlərin monitorinqinin əənəvi üsulları.....	7
1.2. İqtisadi məlumatların təhlili üçün mövcud sistemlərin icmalı.....	12
FƏSİL II. İQTİSADI MƏLUMATLARIN İNTELLEKTUAL TƏHLİLİ MODELLƏRİ VƏ METODLARI.....	18
2.1. Məlumatların intellektual təhlili (DM) faydalı biliklərin tapılması prosesinin əsası kimi (KDD).....	18
2.2. İqtisadi məlumatların intellektual təhlil üsullarının icmalı.....	30
FƏSİL III. İNTELLEKTUAL TƏHLİL ÜSULLARININ TƏTBİQİNİN METODİKİ VƏ PRAKTİK ASPEKTLƏRİ.....	48
3.1. İqtisadi məlumatların intellektual təhlili üçün göstəricilər sisteminin qurulması.....	48
3.2. İqtisadi göstəricilərin intellektual təhlili (DM) üsulları ilə proqnozlaşdırılması.....	65
NƏTİCƏ.....	74
İSTİFADƏ EDİLƏN ƏDƏBİYYAT SİYAHISI.....	75
PE3IOME.....	79
SUMMARY.....	80

GİRİŞ

Mövzunun aktuallığı: Hal hazırda tədqiqat nəzəriyyəsi həm ümumi metodoloji, həm də texnoloji cəhətdən zəngin bir tarixə malikdir. Fərdi təşkilat səviyyəsində bu tədqiqat fəaliyyətlərinin nəticələrini planlaşdırmaq və inkişaf etdirmək üçün tarixən toplanmış təcrübədən istifadə etməyə imkan verir. Tədqiqatın icrası və tədqiqatın idarə olunması sahəsində təcrübənin formalaşdırılması tarixən toplanmış təcrübə məlumatlarından, ekspert biliklərindən və kodlaşdırılmış biliklərdən (məsələn, ontologiya şəklində) əldə edilən məlumatlara əsaslanır.

Çevik və qənaətcil tədqiqat texnologiyalarının inkişafı kontekstində rəşional tədqiqat və idarəetmə qərarlarının qəbul edilməsi üçün faydalı məlumat əldə edən avtomatlaşdırılmış vasitələrin olması vacibdir.

Data Mining, məlumat qranullarının müxtəlif cəhətlərini verilənlər bazasından operativ çıxarmaq üçün əsas vasitələrdən biridir. Bu mənada Data Mining inkişaf etmiş analitiklər, o cümlədən məkan (və ya təsvir) və müvəqqəti (və ya predikativ) analitiklər əsasında tədqiqat fəaliyyətinin avtomatlaşdırılması vasitəsidir.

Xarici mənbələrdə məlumatların intellektual təhlili tez-tez məlumatların əvvəlcədən işlənməsi, təhlili və sonradan emalı da daxil olmaqla daha ümumi bir metodologiya kimi şərh olunur. Məlumatların intellektual təhlili Knowledge Discovery in Databases & Data Mining (KDD & DM) sahəsində biliklərin aşkarlanması anlayışı və ya Data Mining anlayışı ilə əlaqələndirilir. Beləliklə, C. C. Aggarwal-ın Data Mining tərifində bu anlayış məlumat çıxarmaq üçün məlumatların seçilməsi, təmizlənməsi, işlənməsi və təhlilinin avtomatlaşdırılmış metodlarının öyrənilməsi və yaradılması da daxil olmaqla elmi və praktik bir istiqamətdir.

Burada məlumatların intellektual təhlili anlayışının KDD & DM olaraq şərh edilməsinə Data Mining prosesini və vəzifələrini ayrıca vurğulayaraq riayət etmək lazımdır. Tipik olaraq məlumatların təhlili geniş metod və keyfiyyət meyarlarına

malik statistik və ehtimal olunan modellər əsasında aparılır. Lakin obyektlərin və proseslərin ehtimal olunan modeli həmişə həll oluna bilən problemə adekvat ola bilməz. Onda təhlil problemlərini həll etmək üçün təbiət tərəfindən müəyyən olunmuş evristik modellərin və metodların köməyinə müraciət edirlər. Bu cür modellərə süni neyron şəbəkələri, təkamül alqoritmləri və qeyri-səlis sistemlər daxildir. Bütün bunların hamısı məlumatların intellektual təhlili istiqamətinə yönəlmişdir.

Təsviri və predikativ analitika tez-tez yeni məhsul və texnologiyaların inkişafı üçün əsas rəsmi vasitə sayılır. Yalnız yeni deyil, rəqabət qabiliyyətli bir məhsul hazırlamaq üçün analiz mərhələsində formal metodlardan istifadə etməklə ortaya çıxan meylləri öyrənmək lazımdır. Bu iş perspektivli funksiyalara və texnologiyalara, elmi nailiyyətlərə və istifadəçi tələblərinə yönəldilməlidir. Bu cür tədqiqatlar aparmaq üçün faydalı və yeni məlumatlar çıxara bilən alqoritmlərdən istifadə etmək lazımdır. Data mining elmi və praktik bir istiqamət olaraq bu cür alqoritmlərin yaradılması və tətbiqinə yönəldilmişdir. Bu istiqamət toplanmış məlumatların sürətli artması şəraitində fəal inkişaf edir, lakin məlumatların təhlili sistemlərinin tədqiqatı ilə bağlı metodoloji cəhətdən kifayət qədər əhatə olunmur.

Bu mövzunun məqsədi tədqiqat məsələlərinin avtomatlaşdırılmasına diqqət yetirməklə Data Mining-in əsaslarını tanıtmadır. Burada məlumatların yaradılması və əsas anlayışların istiqamətinin inkişaf etdirilməsinin vacibliyi və ehtiyacı məsələləri əhatə olunur. Eyni zamanda Data Mining məsələləri ilə avtomatlaşdırılmış tədqiqat məsələləri arasında qarşılıqlı əlaqəyə, Machine Learning və Knowledge Discovery in Databases və məlumatların təhlili mərhələlərinə baxılır.

Data Mining-in əsas vəzifələrinin formalaşdırılması məlumatların təhlili sistemlərinin tədqiqində çətinliklərə səbəb olur. Riyazi formada tam və ciddiliyi iddia etmədən, burada vahid metodoloji mövqelərdən, təsnifat problemlərinin rəsmi açıqlamaları, anomaliyalar, klasterləşdirmə, proqnozlaşdırma və assosiativ qaydaların çıxarılması aparılır. Rəsmi konseptual analiz əsasında anlayışların axtarışı və lingvistik xülasə yaratmaq kimi nisbətən yeni məlumatların verilməsi vəzifələrinin ifadələri və nümunələri verilmişdir.

Məlumatların intellektual təhlili sistemlərinin inkişafı müəyyən təcrübə qazandığından, onların inkişafı üçün standart bir prosesə keçməyə ehtiyac var. Buna görə KDD & DM sistemlərinin yaradılması sahəsində səkkiz əsas standart metodologiya təqdim olunur.

Mövcud olan müasir tendensiyalardan biri ərazi inkişafının idarəetmə sistemində proqnozlaşdırmanın artan əhəmiyyəti və roludur. Strateji hədəf qərarları müasir proqnoz metodları və vasitələrindən istifadə etməklə əsaslandırılmalı və müəyyən bir ərazinin inkişaf xüsusiyyətlərini nəzərə almalıdır.

Tədqiqat zamanı müasir avtomatlaşdırılmış texnologiyalara əsaslanaraq, daxil olan məlumatların effektiv təhlili və dövlətin sosial-iqtisadi inkişafının proqnozlaşdırılması üçün sistem hazırlanır. Daxil olan məlumatların təhlili və proqnozlaşdırma prosesi təsvir olunur: statistik məlumatların toplanması, məlumat banklarının formalaşdırılması, proqnoz modelinin seçilməsi, məlumatların təhlili və işlənməsi, analitik hesabatların formalaşdırılması. Minimum səhv olan model haqqında məlumatlara əsaslanan sistem ən yaxşı proqnozu seçməyə imkan verir.

Tədqiqatın məqsədi: İqtisadi məlumatların intellektual təhlili üçün göstəricilər sisteminin qurulması, iqtisadi göstəricilərin intellektual təhlili (DM) üsulları ilə proqnozlaşdırılması, məlumatların intellektual təhlili (DM) texnologiyasının araşdırılması və iqtisadi göstəricilərin təhlilindən ibarətdir.

Tədqiqatın predmeti: Tədqiqatın predmet kimi iqtisadi göstəricilər, informasiya texnologiyalarının iqtisadiyyatda rolu, iqtisadi göstəricilərin təhlili və proqnozlaşdırılması üçün Data Mining-in tətbiqi çıxış edir.

Nəzəri və təcrübi əhəmiyyəti: Aparılan tədqiqatın elmi müddəaları və əsas nəticəsindən, tövsiyələr və təkliflərdən, məlumatların intellektual təhlili texnologiyasından istifadə edərək iqtisadi göstəricilərin proqnozlaşdırılıb lazımı biliklərin əldə olunmasıdır.

İnformasiya mənbəyi: Dissertasiya işinin informasiya mənbəyi kimi iqtisadiyyat və informasiya texnologiyaları üzrə xarici və yerli müəlliflərin kitabları və elmi işlərindən istifadə olunmuşdur.

İşin strukturu və həcmi: Tədqiqat işi giriş, 3 fəsil, nəticə və ədəbiyyat siyahısından ibarətdir. Hər fəsil 2 paragrafdan təşkil olunmuşdur. Birinci fəsildə iqtisadi göstəricilərin monitorinqinin metod və sistemləri haqqında məlumat verilib. İkinci fəsildə iqtisadi məlumatların intellektual təhlili modelləri və metodları müəyyən olunub. Üçüncü fəsildə isə intellektual təhlil üsullarının tətbiqinin metodiki və praktik aspektləri müəyyən olunub. İşin sonunda isə nəticə və ədəbiyyat siyahısı verilib.

FƏSİL I. İQTİSADI GÖSTƏRİCİLƏRİN MONİTORİNQİNİN METOD VƏ SİSTEMLƏRİ

1.1. İqtisadi sistemlərin monitorinqinin ənənəvi üsulları

İqtisadi monitorinq sistemi müəssisələrin səmərəli işləməsini təmin etmək üçün zəruridir. Müəssisələrin iqtisadi və maliyyə vəziyyəti haqqında hərtərəfli biliklər investisiya komponentinə əsaslanaraq iqtisadi artımın və inkişafın təmin edilməsinin şərtlərindən biridir.

İqtisadi monitorinqin gəlir bazası az olan bölgələr üçün xüsusi əhəmiyyəti var. Monitorinq daxili iqtisadiyyatın stimullaşdırılmasına əsaslanan mal və xidmət istehsalı üçün effektiv böyümə zonalarının formalaşmasını sürətləndirməklə iqtisadiyyatın real sektorunun sonrakı deqradasiyasını dayandırmağa və gəlir bazasını genişləndirilməyə kömək edə bilər. Monitorinq həm operativ, həm də ortamüddətli rejimdə iqtisadiyyatın ən vacib tendensiylərinin obyektiv qiymətləndirilməsini təmin etməyə imkan verir və bunun nəticəsində elmi və metodoloji əsaslandırılma ilə struktur, investisiya və pul siyasəti alətləri vasitəsilə bu proseslərə təsir göstərmək üçün tədbirlər planı hazırlayır.

Bazar iqtisadiyyatı mühitində, idarəetmə qərarlarının qəbulu zamanı ən vacib yerin qeyri-müəyyən bazar şəraiti, risk və rəqabət qarşısında fəaliyyət variantlarının inkişafına verilməsi səbəbindən bu problemin elmi araşdırması da zəruridir. Bütün bunlar müəssisələrin fəaliyyətini və iqtisadi dinamikasının diaqnozunu əhəmiyyətli dərəcədə yüksəldir.

Monitorinqin mahiyyəti

Monitorinq zəruri statistik hesabatların hazırlanması, sosial-iqtisadi məlumatların toplanıb təhlil olunması, əlavə məlumat və analitik sorğuların (əhali sorğuları və s.) aparılması, inkişaf meyllərinin və müəyyən regional problemlərin vəziyyətinin qiymətləndirilməsi (diaqnostikasi) üçün xüsusi təşkil olunmuş və daim fəaliyyət göstərən sistemdir. [30]

Bu sistemin təşkili, effektivliyi, sistemin bütün elementlərinin necə bağlandığından, bir-birlərinə adekvat olub olmamasından, bu sistemin işlədiyi şərtlərdən asılıdır. Monitoring bir proses olaraq öyrənilən obyekt, onun işlənməsi, sistemləşdirilməsi, təhlili, qiymətləndirilməsi, təsviri, gələcək inkişaf proqnozu haqqında vacib, çoxyönlü məlumat toplanmasının ardıcıl həyata keçirilməsidir.

Monitoringin əsas məqsədi qərarların qəbulu və təhlili üçün informasiya toplamaq, öyrənmək və hazırlamaqdır. Bu da monitoringin informasiya toplamaq və emal etmək sistemi kimi təmin etməli olduğu iki xüsusiyyətə səbəb olur: məlumat proseslərinin hədəf istiqamətləndirilməsi və məlumatların işlənməsinin hər mərhələsində əldə edilən nəticələrin maksimum obyektivliyi.

Monitoring tapşırıqlarına aşağıdakılar daxildir:

- ərazidə gedən proseslərin gedişi barədə müşahidələrin təşkili, obyektiv və etibarlı informasiyanın əldə edilməsi;
- alınan məlumatların qiymətləndirilməsi və sistem təhlili;
- idarəetmə orqanlarına, müəssisə, idarə, təşkilatlara, tabeliyindən və mülkiyyət formasından asılı olmayaraq, vətəndaşları monitoring zamanı əldə edilmiş məlumatlarla təmin etmək;
- sosial və iqtisadi təhlükələrə səbəb olan amillərin gələcək və indiki zaman üçün müəyyən edilməsi;
- sosial-iqtisadi vəziyyətin inkişafı üçün proqnozların hazırlanması;
- mənfi istiqamətləri aradan qaldırmaq, müsbət meylləri dəstəkləmək və onları regional hakimiyyət orqanlarına çatdırmaq üçün tövsiyələrin hazırlanması.

Monitoring işləri beş mərhələdə aparıla bilər:

Birinci mərhələdə statistika, müəssisə və təşkilatların məlumatları və ekspert rəy sorğularından istifadə etməklə həyata keçirildiyi ehtimal olunan məlumatlar toplanır.

İkinci mərhələ, yaradılan məlumat bazasından istifadə edərək müvafiq nəticə hazırlamaq məqsədi ilə araşdırmanın başa çatdığı tarixdə və gələcək üçün araşdırılan obyektin cari vəziyyətinin diaqnostikasıdır.

Monitorinqin üçüncü mərhələsində araşdırılan obyektin cari vəziyyətinin diaqnostikasının nəticələrinə əsasən qısa, orta və ya uzun müddətə mümkün inkişaf variantları hazırlanır. Bir qayda olaraq, bir neçə alternativ variant tərtib olunur, bunlardan ən əlverişlisi tövsiyələrin hazırlanması və qərarların qəbulu hesab olunur.

Dördüncü və beşinci mərhələlərdə tövsiyələr verilir.

Monitorinq - iqtisadi təhlilin qeyri-rəsmi metodu kimi

Hal-hazırda iqtisadi təhlil metod və texnikalarının müxtəlif təsnifatları mövcuddur. Bütün təsnifatlar fərqli xüsusiyyətlərə əsaslanır. Ən məlumatlı üsul prosedurların təsviridir. Bu məntiqə əsaslanaraq bütün analitik metodları qeyri-rəsmi (məntiqi) və rəsmi metoda (riyazi) bölmək olar.

Qeyri-rəsmi metodlar, analitik asılılıqların köməyi olmadan prosedurların məntiqi səviyyədə təsvir edilməsinə əsaslanır. Burada, mühüm rolu təcrübə, analitikin intuisiyası oynayır.

Bu cür üsullara aşağıdakılar daxildir:

- göstəricilər sisteminin işlənməsi;
- müqayisə metodu;
- analitik cədvəllərin qurulması;
- təfərrüatları qəbul etmək;
- ekspert qiymətləndirmə metodu;
- situasiya təhlili və proqnozlaşdırma metodları;
- ssenari metodu;
- simulyasiya modelləşdirmə;
- monitorinq.

Formallaşdırılmış üsullar əvvəlcədən müəyyən edilmiş ciddi asılılıqlara və qaydalara əsaslanır.

Əksər hallarda, əsasən qeyri-rəsmi metodlardan, eləcə də iqtisadi analiz və statistikanın klassik metodlarından istifadə edir. Bu gün iqtisadi və maliyyə fəaliyyətin qiymətləndirilməsində yeni bir metod olan monitorinq, hələdə dövlətin iqtisadi nəzəriyyəsi və praktikasında lazımi şəkildə istifadə edilməmişdir. Monitorinq

iqtisadi metod kimi (tədqiqatlar, ekspert qiymətləndirmələri, iqtisadi qiymətləndirmələr və s.) göstəricilər sistemi və onun həyata keçirilmə məqsədlərinə, tədqiqat subyektlərinin və obyektlərinin inkişaf şərtlərinə əsasən anlayışın müəyyənləşdirilməsi üçün əsas kimi qəbul edilmişdir. Monitoring müəyyən bir obyektin davranışını araşdırma, qiymətləndirmə, təhlil və proqnozlaşdırmağa imkan verən elmi bir metod olaraq başa düşür.

Monitoring müəssisənin maliyyə-təsərrüfat fəaliyyətini qiymətləndirmə metodu olaraq bu gün onun inkişaf istiqamətini müəyyənləşdirir və təkcə ekspert qiymətləndirmə metodlarına əsaslanmır eyni zamanda iqtisadi proseslərin dinamikasını görməyə, bunun arxasında duran amilləri qiymətləndirməyə imkan verən metodlar, üsullar, alətlər toplusunu da əhatə etməlidir eyni zamanda müasir iqtisadi münasibətlər şəraitində bütövlükdə müəssisənin inkişaf xüsusiyyətlərinə onların təsir səviyyəsini müəyyənləşdirməlidir.

Müəssisənin maliyyə-təsərrüfat vəziyyətinin monitorinqi

idarəetmə aparatının informasiya təminatı probleminin həllində ən perspektivli yanaşmalardan biri kimi müəssisənin maliyyə vəziyyətinin monitorinqinin təşkili və həyata keçirilməsi çıxış edir. Müəssisənin idarə edilməsində bir müəssisənin maliyyə vəziyyətinin müntəzəm izlənilməsinin rolunu qiymətləndirmək çətindir. Mikro və makro səviyyədə iqtisadi vəziyyətdəki dəyişikliklərin operativ və müstəqil qiymətləndirilməsini əldə etmək, iqtisadi proseslərin təhlili və proqnozlaşdırılması imkanını verir. Müəssisənin maliyyə vəziyyətinin obyektiv monitorinqi kritik vəziyyətləri və mənfi meyilləri vaxtında müəyyənləşdirməyə və vəziyyəti vaxtında düzəltməyə imkan verir, yəni hər zaman müəssisəni ayaqda saxlamağa kömək edir.

Maliyyə müəssisəsinin monitorinqi daim dəyişən maliyyə bazarı şəraitində maliyyə fəaliyyətinin ən vacib cari nəticələrini daim izləmək üçün mexanizm olaraq təyin olunur.

Müəssisənin maliyyə vəziyyətinin monitorinqi sisteminin qurulmasının məqsədləri aşağıdakılardan ibarətdir:

- müəssisənin maliyyə vəziyyətindəki mənfi meyillərin vaxtında müəyyənləşdirilməsi və təhlili;
- böhran vəziyyətlərinin qarşısının alınması;
- müxtəlif səviyyələrdə idarəetmə sahəsində operativ və uzunmüddətli idarəetmə qərarlarının hazırlanması üçün idarəetmə orqanlarına müvafiq, etibarlı və dolğun məlumat vermək.

Hər hansı müəssisənin fəaliyyəti yalnız istehsalın təşkili və onu lazımı mənbələrlə təmin etməklə bitmir, eyni zamanda müəssisənin cari fəaliyyətini daim izləmək və planlaşdırılan nəticələrə nail olmaq üçün idarəetmə qərarlarına düzəliş etmək lazım gəlir. Bütün bunlar müəssisənin əsas iqtisadi və maliyyə nəticələrinin əvvəlcədən hesablanmış və planlaşdırılmış göstəricilərlə daimi müqayisə edilməsini zəruri edir.

Zaman vahidi üçün qəbul edilən və təyin olunan ən vacib göstəricilər sırasına aşağıdakılar daxildir:

- malların və göstərilən xidmətlərin satışından əldə olunan gəlir;
- müəssisənin istehsal fəaliyyətini həyata keçirmək üçün ümumi xərclər;
- müəssisənin vergiyə cəlb edilməsindən əvvəl mənfəət;
- istehsalın rentabelliyi;
- ümumi gəlir;
- işçilərin əmək haqqı səviyyəsi;
- şirkətin cari hesabındakı vəsaitlər;
- müəssisənin kreditör borcları;
- debitor borcları.

Ən vacib göstəricilərin planlaşdırılmış dəyərləri - bu şirkətin, idarə tərəfindən nəzarət edilən çoxşaxəli idarəetmə fəaliyyətinin yalnız bir sahəsidir. Onun həyata keçirilməsini təmin etmək üçün daim müşahidə etmək lazımdır ki, həqiqi nəticələr planlaşdırılan nəticələrdən az olmasın yəni, bu o deməkdir ki daimi monitoring tələb olunur.

Cəmiyyətin faktiki cari vəziyyəti barədə mövcud məlumatlar rəhbərlərin və rəhbərlik etdikləri xidmətlərin diqqət mərkəzində olmalıdır. Beləliklə, zəruri hallarda

təcili olaraq müəyyən edilmiş nöqsanların aradan qaldırılması üçün onların konkret tədbirlər hazırlayıb təklif etmələri məqsədə uyğundur.

1.2. İqtisadi məlumatların təhlili üçün mövcud sistemlərin icmalı

Avtomatik məlumat toplama vasitələri, məlumatların saxlanması üçün informasiya texnologiyalarının geniş tətbiqi, elektron sənəd idarəetməsi, şəbəkə texnologiyaları - bütün bunlar, saxlanılan məlumatların strukturunun, həcmnin və mürəkkəbliyinin artmasına səbəb olur. Bununla əlaqədar olaraq, daha çox istifadə və faydalı məlumat çıxarmaq üçün böyük həcmli, mürəkkəb quruluşlu avtomatlaşdırılmış məlumatların intellektual təhlili vasitələrinin inkişafına ehtiyac var. Bu məqsədlər üçün statistika, məlumat nəzəriyyəsi kimi müxtəlif elmi sahələrə aid problemlərin həlli üçün metodlar hazırlanmışdır. Son on illiklərdə bu tədqiqat sahələri Data Mining, Machine Learning və Knowledge Discovery in Databases sayəsində genişləndirilib.

Data Mining, Machine Learning və Knowledge Discovery in Databases

Ədəbiyyatda, məlumatların intellektual təhlili sahəsində fəal inkişaf edən üç qrup araşdırılır: Data Mining, Machine Learning ı Knowledge Discovery in Databases (KDD). Bəzən bu sahələr ekvivalent hesab olunur, çünki məlumat modellərini, xüsusiyyətlərini və asılılıqlarını araşdırmağa və qurmağa yönəldilmişdir.

Hal-hazırda belə hesab olunurki KDD, Data Mining –ə bir nüvə kimi daxil olur və bunun nəticəsində Machine Learning-in metodlarından istifadə edir. Data Mining

saxlanan məlumatları təhlil etmək üçün hazırlanmış informasiya texnologiyalarının təbii inkişafının bir hissəsidir.

Data Mining termini mürəkkəb problemlərin avtomatlaşdırılmış həllinə görə bir sıra metodlar dəstinin təyin edilməsi üçün böyük məlumat bazalarından faydalı məlumatlar çıxarmaq məqsədi ilə meydana gəlmişdir. Sonralar isə Data Mining məlumatlarda xassələr, qruplaşmalar, asılılıqlar şəklində və bu məlumatın sonrakı tətbiqi, iqtisadi və ya digər faydalar üçün tətbiq edilmişdir. Bu məlumatlar çox vaxt bir domen nümunəsi olaraq adlandırılır, və o mövcud sahədəki mövcud vəziyyəti eyni zamanda mümkün dəyişiklikləri xarakterizə edir.

Beləliklə, məlumatların təhlili istiqamətinin qurulması və inkişafı üçün aktualıq və ehtiyac "Big Data" problemi ilə əlaqədardır. Bununla əlaqədar olaraq böyük məlumatlardan faydalı məlumatların avtomatlaşdırılmış çıxarılması üçün müxtəlif elmi paradigmlərin metodlarını inkişaf etdirmək zərurəti yaranmışdır.

Şəkil 1.-də bu anlayışlar arasındakı əlaqə, metod və modellərin digər sinifləri ilə əlaqə göstərilmişdir. Statistik təhlil, verilənlər bazası, məlumat nəzəriyyəsi, optimallaşdırma metodlarına əsaslanaraq KDD & DM, böyük məlumatlar üçün "intellektual" problemlərin həllinin daha yüksək avtomatlaşdırılması səbəbindən daha qısa müddətdə daha dəqiq nəticələr əldə etmək məqsədi daşıyır.



Şəkil 1. Məlumatların çıxarılması istiqamətini və aralarında Data Mining-in yerini təşkil edən metodlar toplusu

Buna görə də KDD & DM prosesinin məzmununu məlumatların təhlili sistemlərinin tədqiqatı üçün yararlı bir sistem metodologiyası kimi təqdim etmək lazımdır.

KDD & DM metodologiyası faydalı məlumat çıxarmaq üçün yerinə yetirilməli olan hərəkətlərin ardıcılığını təsvir edir. Bu addımlar ardıcılığı mövzu sahəsindən asılı deyil, lakin mütəxəssis biliklərinin istifadəsini də istisna etmir.

KDD & DM metodologiyasına görə məlumatların intellektual təhlili sistemlərinin tədqiqatı üçün 5 konseptual mərhələ vardır:

1. Həll ediləcək problemin ifadəsi:

a. mövzu sahəsi, təhlil məqsədləri, tətbiqetmənin iş vəzifələrini anlamaq, istifadə halları haqqında lazımlı biliklərin müəyyən edilməsi və formalaşdırılması;

b. problemin həll yollarını irəli sürmək və lazımi dəyişənləri təyin etmək;

c. həll ediləcək Data Mining probleminin növünün seçimi (təsnifat, proqnozlaşdırma, klasterləşdirmə, anomaliyaların axtarışı (istisnalar), assosiativ qaydaların axtarışı və s.);

d. həll ediləcək problemin formalaşdırılması. Problem məlumatların intellektual təhlili sisteminin tədqiqatçısı və mövzu mütəxəssisi arasında sıx əməkdaşlıq nəticəsində formalaşmışdır.

2. Təhlil üçün məlumatların yaradılması və istifadəsi:

a. xam məlumatların axtarışı (və ya seçimi) və toplanması (konsolidasiya) alt sisteminin tətbiqi ilə mümkündür. Bu müddətdə iki yanaşma mövcuddur: birincisi mütəxəssisin nəzarəti altında məlumatların toplanması və ikincisi məlumatların təsadüfi seçilməsi;

b. məlumatların əvvəlcədən işlənməsi (normallaşdırma, diskretizasiya, itirilmiş dəyərlərin işlənməsi, artefaktların silinməsi, homogenliyin yoxlanılması). Çox vaxt

bu mərhələdə lazımsız faktorlar aşkarlanır və çıxarılır, məlumatlar çevrilir və ölçülür, xarakterik xüsusiyyətlər və ya xüsusiyyətlər məlumatlardan çıxarılır;

c. ölçü azaldılması, əhəmiyyətli xüsusiyyətlərin seçilməsi, inteqral göstəricilərin və invariantların hesablanması.

3. Data Mining alqoritminin seçilməsi, tətbiqi və qiymətləndirilməsi:

a. Data Mining alqritmi üçün məhdudiyyət və tələblərin dəqiqlik, ölçü, təfsiretmə, qurulma sürəti, əldə edilən modellərin tətbiqi, mənbə məlumatlarının növü kimi müəyyənləşdirilməsi;

b. ilk mərhələdə seçilmiş tapşırıqların həlli üçün bir və ya daha çox alqoritm seçimi;

c. Data Mining əsasında problemin həllini qiymətləndirmək üçün meyarlar seçimi;

d. seçilən Data Mining alqoritminin tətbiqi, sınılanması və qiymətləndirilməsi (vizuallaşdırma, təsvir, boşluğun aradan qaldırılması, dəqiqliyin qiymətləndirilməsi, modellərin etibarlılığı və s.).

4. Faydalı məlumatların çıxarılması üçün seçilmiş Data Mining alqoritmlərinin tətbiqi:

a. hasil olunan məlumatın faydalılığının qiymətləndirilməsi;

b. bir çox dəyişən arasında analiz nəticələrinin yığılması.

5. Təhlil nəticələrinin şərh edilməsi və təqdim edilməsi:

a. Data Mining nəticələrinin mətnli təsvirini yaratmaq üçün bir vasitə kimi dil xülasəsi;

b. nəticələri verilənlər bazasında və bilik bazalarında saxlamaq.

Tapşırıq və nəticələrin növləri

Analitiklər və tədqiqatçılar üçün Data Mining problemlərinin həlli prosesində çıxarılan məlumatlar aşağıdakı növlərə bölünə bilər:

- Faydalı. Faydalı olan məlumat qərarları doğru vermək üçün istifadə edilə bilər.
- Təcrübəsiz. Məlumat əvvəllər məlum idi və əvvəl məlum olan informasiya qərarlarının yerinə yetirilməsini yoxlamaq üçün istifadə edilə bilər.

• Anlaşılmaz. Çıxarılan məlumatlar əvvəllər məlum deyildi, lakin bu məntiqi izah edilə bilməz. Bu cür məlumatlar bir tərəfdən tamamilə yeni və gələcəkdə faydalı olacaq nümunələri, digər tərəfdən tamamilə yararsız və ziddiyyətli qranulları ehtiva edə bilər. Bu da məlumatlarda xüsusi nümunələrin tapılmadığını göstərir. Bu o deməkdir ki, ya bu məlumatlar ümumiyyətlə yoxdur, ya da onların aşkarlanması üçün fərqli bir metod seçmək lazımdır. Aydınlaşdırmaq üçün isə əlavə bir analiz tələb olunur.

"Faydalı məlumat" termini məlumatın tədqiqat fəaliyyətinin subyekti üçün gələcəkdə müəyyən mənfəət və ya qazanc əldə etməsi baxımından əhəmiyyətini ifadə edir. Eyni zamanda subyekt rolunda müəyyən bir təşkilat (və ya tədqiqat komandası) və ayrıca bir tədqiqatçı çıxış edə bilər. Bu mənada, məlumatın faydalılığı dövlətin qiymətləndirilməsindən, inkişaf istiqamətinin seçilməsindən və mövzunun elan edilmiş məqsədlərindən, yəni məkan-müvəqqəti proqnozlardan və əldə edilmiş nəticələrdən asılı olaraq nisbi xarakter daşıyır. Şübhə yoxdur ki, məlumatın faydalılığı səviyyəsinin təfsiri ekspert qiymətləndirilmələri əsasında ifadə olunur və yuxarıda göstərilən xüsusiyyətlərin intellektual təhlili metodlarına əsaslanır.

Data Mining istiqamətində faydalı məlumatları çıxarmaq üçün təsviri və proqnostik (proqnozlaşdırıcı) təhlil tapşırıqları həll olunur. Təsviri tapşırıqların birinci sinfinə aşağıdakılar daxildir:

- qruplaşmaların yaradılması vəzifələri (bölmə, klasterləşdirmə, formal konsepsiyaların təhlili);
- məlumatların xassələrini çıxartmaq vəzifələri (təsnifat, linqvistik, xülasə, anomaliya axtarışı). İkinci sinif predikativ analizin vəzifələrini ehtiva edir.
- məkan və müvəqqəti asılılıqların qurulması vəzifələri (proqnozlaşdırma, korrelyasiya təhlili, reqressiya təhlili, assosiativ qaydaların axtarışı;
- atributlar (təsnifat) məkanında asılılıqların qurulması vəzifələri.

Tədqiqat fəaliyyətinin tapşırıqlarının Data Mining vəzifələri ilə əlaqəsi

Avtomatlaşdırılmış tədqiqat sistemləri faydalı məlumatların çıxarılması mürəkkəbliyini və tədqiqat müddətini azaltmaqla mühəndislərin səmərəliliyini

artırmaq məqsədi daşıyır. Adətən bu, tədqiqat fəaliyyətinin aşağıdakı vəzifələrini həll etməklə əldə edilir:

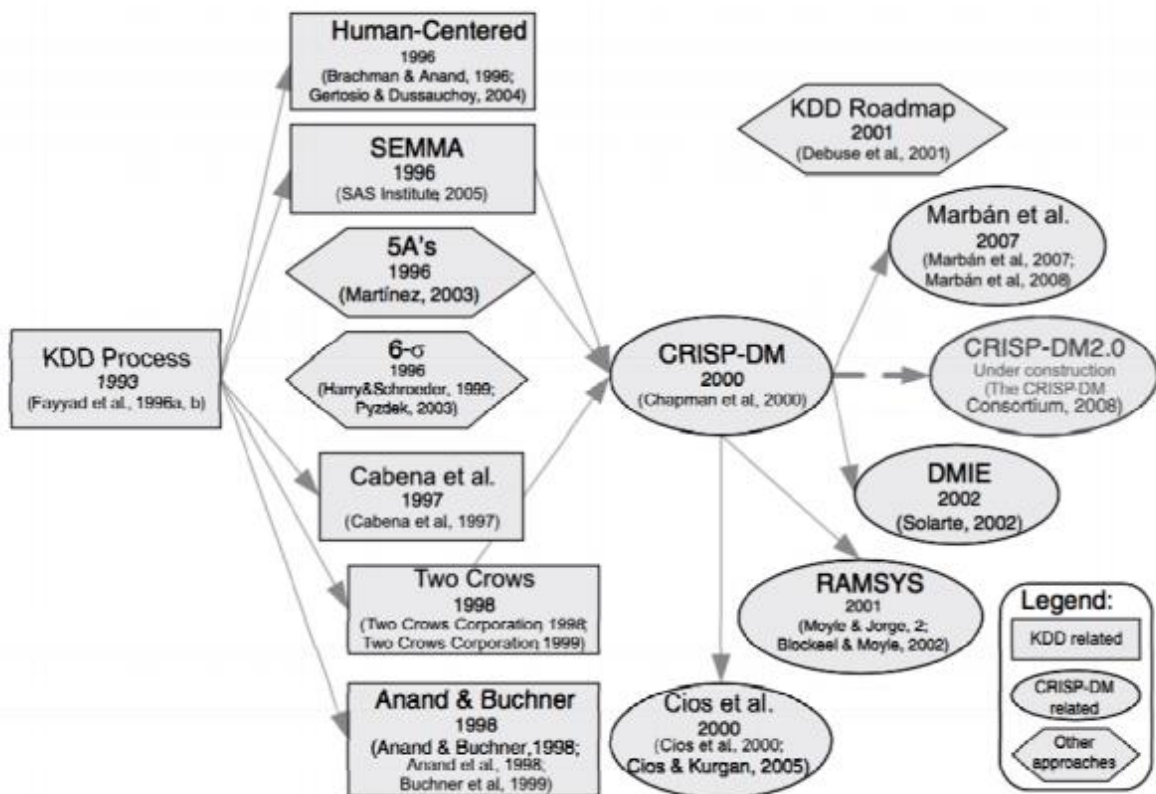
1. sənədləşmə işlərinin (sənədlərin) avtomatlaşdırılması;
2. qərar qəbuletmə prosesinin (dəstək və qərar qəbulu) avtomatlaşdırılması;
3. tədqiqat qərarlarının, məlumatların və işlənmələrin təkrar istifadəsi üçün tədqiqat qərarlarının və proseslərinin birləşdirilməsi (analoqlardan istifadə);
4. strateji tədqiqatların yaradılması (tədqiqat vəziyyətinin öyrənilməsi, əlaqələrin və asılılıqların modelləşdirilməsi);
5. sahə testlərinin və prototiplərin riyazi modelləşdirmə ilə əvəz edilməsi;
6. tədqiqat idarəetmə keyfiyyətinin yaxşılaşdırılması (uyğunsuzluq və yenilik axtarışı, əlaqələrin axtarışı);
7. variantlı tədqiqat və optimallaşdırma metodlarının tətbiqi (əlaqələrin axtarışı, qiymətləndirmə və qərar qəbulu).

FƏSİL II. İQTİSADI MƏLUMATLARIN İNTELLEKTUAL TƏHLİLİ MODELLƏRİ VƏ METODLARI

2.1. Məlumatların intellektual təhlili (DM) faydalı biliklərin tapılması prosesinin əsası kimi (KDD)

Bu fəsildə, verilmiş məlumatların yaradılması prosesini (KDD & DM) müəyyənləşdirən standartlaşdırılmış modelləri (metodologiyaları) nəzərdən keçiririk. Təqdim olunan metodologiyalar əsas Data Mining (DM) prosesinə əlavə olaraq əvvəlcədən işlənmədən və sonradan işlənmədən ibarət olan KDD prosesini mahiyyətə təsvir edir. İcra nöqteyi-nəzərindən ayrı komponentlər hesab olunan KDD & DM prosesinin mərhələləri məlumatların yaradılması sistemlərinin tədqiqatçılarına avtomatlaşdırma dərəcələrini müəyyənləşdirməyə və müxtəlif memarlıq konstruksiyaqlarını qurmağa imkan verir.

Şəkil 2.-də Data Mining sistemlərini inkişaf etdirmək üçün faydalı olan standartlaşdırılmış metodologiyaların təkamülü göstərilmişdir. Şəkil 2. göstərir ki, KDD analiz metodologiyasının başlanğıc nöqtəsidir və CRISP-DM metodologiyası sonrakı yanaşmaların çoxunun əsaslandığı əsas metodologiya kimi təqdim edilmişdir. Göstərilən KDnuggets sorğularının nəticələrinə görə sorğuda iştirak edən şirkətlərin 42% -i CRISP-DM metodologiyasından, 10% -i SEMMA metodologiyasından, 6% -i öz təşkilatının metodologiyasından, 28% -i öz metodologiyasından, 6% -i digər metodologiyadan istifadə edir. Respondentlərin 7% -i heç bir metodologiyadan istifadə etmir. [15]



Şəkil 2. Data Mining sistemlərinin inkişafında standart metodologiyaların təkamülü

SEMMA metodologiyası

SEMMA metodologiyası SAS Data Mining Solution-də tətbiq olunur (1996). Qısaltması aşağıda verilmiş sözlərdən irəli gəlir:

Sample (S) - Nümunə ("Məlumat seçimi", yəni bir seçim yaratmaq);

Explore (E) - Tədqiq et ("Verilənlərdəki əlaqələrin araşdırılması");

Modify (M) - Dəyişdirin ("Məlumatın dəyişdirilməsi");

Model (M) - Model ("Modelləşdirmə asılılığı");

Assess (A) - Qiymətləndirmə ("Yaranan modellərin və nəticələrin qiymətləndirilməsi").

KDD & DM məlumat analizinin SEMMA metodologiyasına uyğun həyata keçirilməsi prosesi Şəkil 3.-də təsvir edilmişdir.

Nümunə götürmə	Tədqiqat	Modifikasiya	Modelləşdirmə	Qiymətləndirmə
- Məlumat mənbəyinə qoşulma - Ümumi əhalidən nümunə götürmə	- Vizual tədqiqat - Assosiasiya siyaların müəyyənləşdirilməsi - İnteraktiv statistik təhlil	- Çoxluq təhlili - Çevrilmə, süzülmə, qarışıqlıq	- Neyron şəbəkələri - Qərarlar ağacı - İstifadəçi tərəfindən müəyyən edilmiş modellər - reqressiya	- Model müqayisəsi - Model hesablamaları

Şəkil 3. SEMMA metodologiyasının mərhələləri

SEMMA metodologiyası Data Mining sistemlərinin yaradılması üçün bütün zəruri məlumatların işlənməsi və təhlili proseslərinin həyata keçirilməsini nəzərdə tutur. SEMMA yanaşması prosesin quruluşunu və hər bir addımı dəstəkləyən vasitələrin məntiqi təşkilini özündə birləşdirir. SEMMA metodologiyası statistik tədqiqat və vizuallaşdırma metodlarının tətbiqini asanlaşdırır, ən əhəmiyyətli dəyişənləri seçməyə və dəyişdirməyə, nəticələri proqnozlaşdırmaq üçün bu dəyişənlərlə modellər yaratmağa, modelin düzgünlüyünü təsdiq etməyə və modeli yerləşdirməyə imkan verir. Bu metodologiya heç bir ciddi qaydalar tətbiq etmir. SEMMA metodologiyasından istifadə nəticəsində mütəxəssis, tədqiqat konsepsiyasının qurulması, həyata keçirilməsi, həmçinin KDD & DM sisteminin tədqiqat nəticələrinin qiymətləndirilməsi üçün elmi metodlara sahib ola bilər.

CRISP-DM metodologiyası

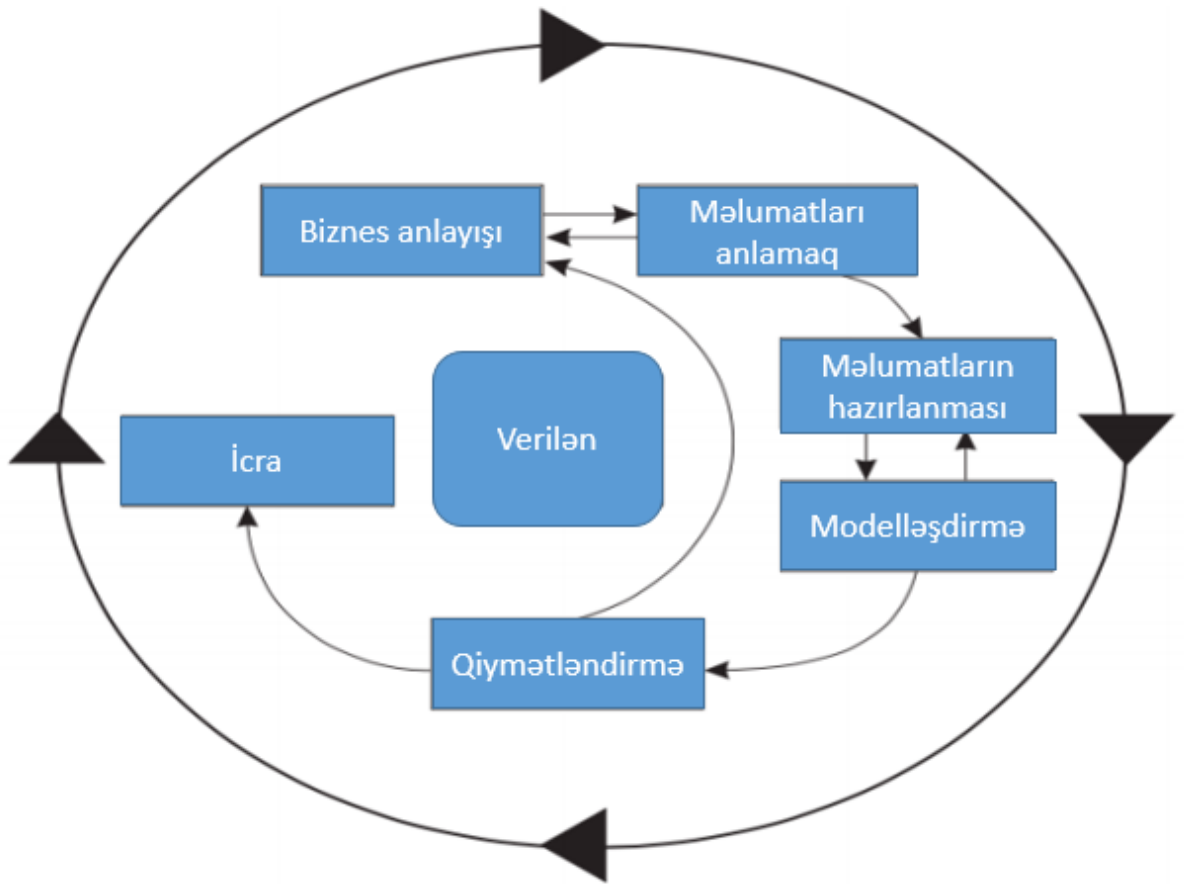
CRISP-DM (The Cross Industrial Standard Process for Data Mining) metodologiyası (Chapman et al., 2000), sənədli və sərbəst paylanmış bir model təqdim edir. Təşkilatların Data Mining metodlarından istifadənin faydalarını maksimum dərəcədə artırması üçün həyata keçirilə biləcək əsas mərhələləri təsvir edir. Bu metodologiya ən populyar və ümumi metodologiyadır.

CRISP-DM üçün ən vacib müvəffəqiyyət amilli platformanın müstəqilliyi və müxtəlif tətbiq sahələrinə uyğunlaşmasıdır.

CRISP-DM standartına görə, KDD & DM bir çox döngəsi və əks əlaqəsi olan bir prosesdir.

Məlumatların yaradılması sistemi layihəsinin həyat dövrü (Şəkil 4) altı mərhələdən ibarətdir. Üstəlik, addımların ardıcılığı ciddi deyil. Oxlar mərhələlər arasındakı ən vacib və tez-tez baş verən əlaqələri göstərir. Şəkil 4.-dəki xarici dairə tsiklik məlumatların hasilatını göstərir.

Proses zamanı alınan dəqiqləşdirmələr digər daha konkret məsələlərə səbəb ola bilər. Sonrakı məlumatların əldə edilməsi, əvvəlkilərdən faydalanmaqla alınır.



Şəkil 4. CRISP-DM standartına uyğun olaraq KDD&DM mərhələləri

CRISP-DM standartına uyğun olaraq KDD&DM mərhələlərini nəzərdən keçirək:

1. Biznes anlayışı:

Bu ilkin mərhələ tədqiqatın məqsədini və tələblərini bir iş nöqtəyi-nəzərindən araşdırmağa və sonra bu məlumatları Data Mining tapşırığının rəsmi açıqlamasına çevirməyə, habelə hədəflərə çatmağa yönəlmiş ilkin plan hazırlamağa həsr edilmişdir.

2. Məlumatları anlamaq:

Məlumatları anlamaq mərhələsi, ilkin məlumatların toplanması, məlumatlarla tanışlığa keçid və məlumatların keyfiyyət problemlərini müəyyənləşdirməklə başlayır. Verilənlərin quruluşunu başa düşmək, gizli nümunələrin sonrakı təhlili və fərziyyələrin formalaşması üçün maraqlı altçoxluqlar tapmaq lazımdır.

3. Məlumatların hazırlanması:

Məlumatların hazırlanması mərhələsi, ilkin məlumatlar toplusundan istifadə edərək son məlumatları müəyyən etmək üçün istifadə olunan bütün işləri əhatə edir. Bu mərhələdə cədvəllər bir sıra qeydlər və atributlar (MD ($n \times d$) verilənlər bazası), həmçinin modelləşdirmə üçün zəruri dəyişikliklər və məlumatların təmizlənməsi ilə yaradılır.

4. Modelləşdirmə:

Bu fazada, ilk mərhələdə müəyyənləşdirilmiş problemin rəsmi açıqlaması və onların tətbiqinə görə Data Mining metodlarının seçimi mövcuddur. Bundan əlavə, bu mərhələdə Data Mining modellərinin uydurma parametrləri, təlim və test halları haqqında faydalı məlumat əldə etmək üçün istifadə olunur.

5. Nəticələrin qiymətləndirilməsi:

Data Mining modelinin son yerləşdirilməsinə başlamazdan əvvəl bu modeli daha dərinlən qiymətləndirmək, modelin qurulmasında iştirak edən bütün addımları qiymətləndirmək, həmçinin biznes tapşırıqlarının həlli üçün faydalı məlumat çıxartmaq problemini həll edib-etmədiyini anlamaq vacibdir. Bu mərhələnin sonunda məlumatların əldə edilməsinin nəticələrindən istifadə edilməsi barədə qərar qəbul edilir.

6. İcra:

Əgər Data Mining modeli formalaşıbsa, bu heç də tədqiqatın başa çatması demək deyil. Alınan məlumatlar müştərinin şərh edə və işlərində istifadə edə biləcəyi şəkildə təqdim edilməlidir. Tələblərdən asılı olaraq yerləşdirmə mərhələsi ya sadə (sadə hesabat yaratmaq) ya da daha mürəkkəb (təkrar məlumatların çıxarılmasını tələb edə bilər) ola bilər.

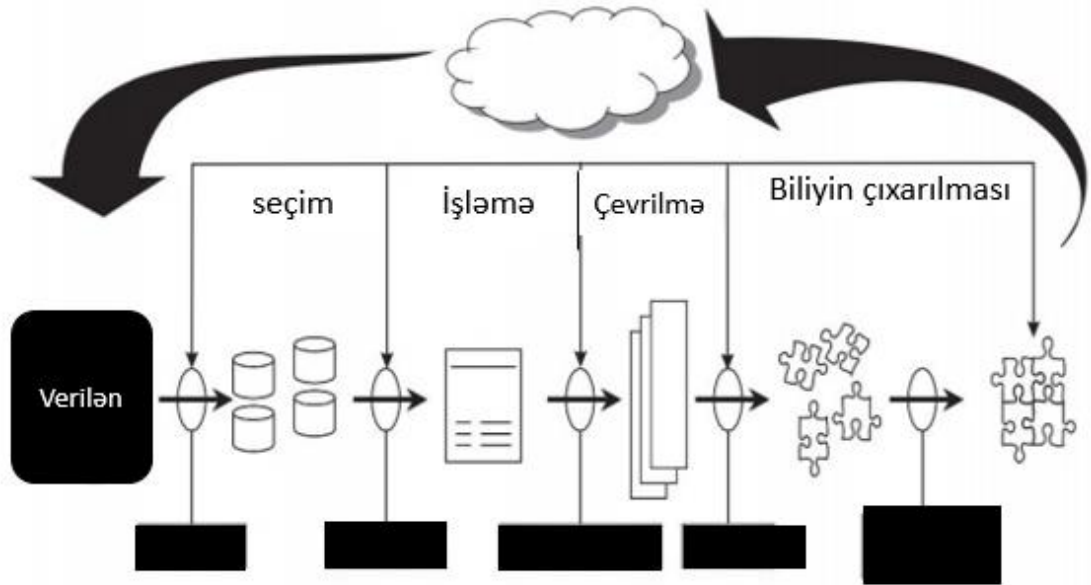
Cabena metodologiyası

Cabena metodologiyasında (Cabena et al., 1997), məlumatların işlənməsi mühüm iş qərarlarının qəbul edilməsində istifadə üçün əvvəllər məlum olmayan, etibarlı və faydalı məlumatları böyük verilənlər bazalarından çıxarmaq prosesi olaraq təyin olunur.

Cabena metodologiyasının addımları Şəkil 5.-də verilmişdir. Bu metodologiyada aşağıdakı 4 addım vurğulanmışdır:

- seçim,
- öncədən işləmə,
- çevrilmə,
- faydalı məlumatların, yeni biliyin çıxarılması.

Son mərhələ, hasil olunan məlumatların biznes üçün faydalı bilik baxımından təhlili və şərhidir. Burada əldə olunan biliklər saxlanılır və sonradan təhlil üçün istifadə olunur. Baxılan metodologiyanın üstünlüyü isə KDD&DM sistemlərinin tədqiqat üçün əlverişli, sadə və anlaşılan şəkildə olmasıdır.



Şəkil 5. Cabena metodologiyasının mərhələləri

Two Crows metodologiyası

Oxşar addımlar üçün fərqli adlar istifadə olunmasına baxmayaraq Two Crow metodologiyası (1998) KDD&DM prosesini təsvir etməyə çox yaxındır.

Two Crow metodologiyasının addımları Şəkil 6.-da verilmişdir. Cabena metodologiyasından fərqli olaraq Two Crow metodologiyası yeddi mərhələni ayırır və gerilmələri son mərhələdən birincisinə deyil, aralıq mərhələlər arasında göstərir ki, bu da daha çox məlumat almağa imkan verir. Beləliklə, Two Crow metodologiyasının əsas addımları aşağıdakılardan ibarətdir:

- biznes tapşırığını müəyyənləşdirmək;
- Data Mining üçün verilənlər bazası yaratmaq;
- məlumatları araşdırmaq;
- modelləşdirmə üçün məlumat hazırlamaq;
- Data Mining üçün bir model yaratmaq;
- modeli qiymətləndirmək;
- modeli və nəticələrini tətbiq etmək.



Şəkil 6. Two Crow metodologiyasının mərhələləri

Digər tərəfdən, bu metodologiyanın qrafik təsviri addımlar, mənbənin növləri və nəticədə əldə olunan məlumatlar arasında əlaqə yaratmır, həmçinin təhlil nəticələrinin saxlanma imkanını göstərmir.

RAMSYS metodologiyası

RAMSYS metodologiyası (Rapid collaborative data mining system, 2001.) problemlərin metodoloji həlli ahənginə, bilik mübadiləsinə və kommunikasiya rahatlığına istiqamətlənmişdir. RAMSYS metodologiyası qiymətləndirmə mərhələsində əldə edilmiş ən yaxşı modellərin məcmusu kimi son Data Mining modelini təqdim etməkdən ibarət olan yeni bir vəzifəni əhatə edir.

Bu metodologiyanın qrafik təsvirindəki fərq əslində mikro mərhələlərin adı olan KDD & DM prosesinin hər mərhələsi üçün tapşırıqların tərtib edilməsidir.

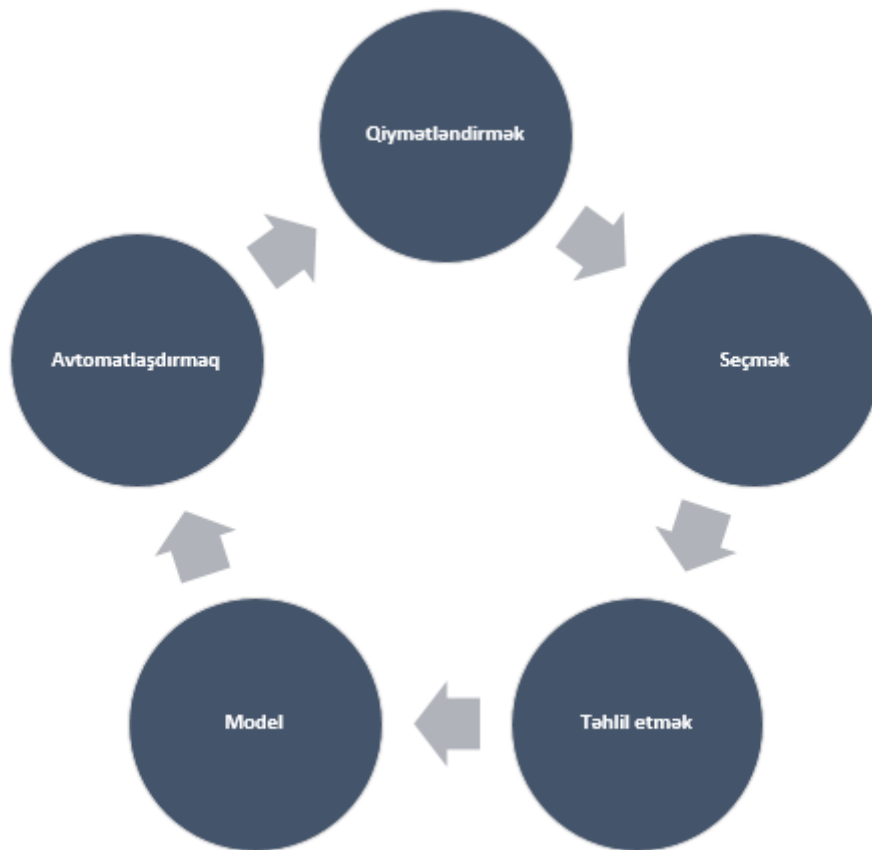
Five A's metodologiyası

Five A's metodologiyası SPSS (2007) şirkəti tərəfindən təklif olunmuşdur. Bu metodologiyanın adı, Data Mining tsiklik prosesinin mərhələlərinin ilk hərfləri ilə formalaşır:

- qiymətləndirmək;
- seçmək;
- təhlil etmək;
- model;
- avtomatlaşdırmaq.

Bu metodologiyanın müsbət tərəfi, məlumatların təhlili sahəsində mütəxəssis olmayan istifadəçilərin əvvəllər əldə edilmiş modelləri yeni məlumatlara tətbiq edə bilməsi üçün məlumatların yaradılması prosesinin avtomatlaşdırılması ideyasıdır.

Five A's metodologiyasının mənfi tərəfi odur ki, qurulmuş modeli tətbiq etmək və ya bilik tapmaq üçün başqa alternativ yollar yaratmır. Digər vacib bir mənfi cəhət odur ki, modelin keyfiyyətini anlamaq və yoxlamaq, habelə tədqiqatın inkişafı zamanı mümkün problemlərin qarşısını almaq üçün metodologiyada CRISP-DM-də vacib sayılan məlumatların başa düşülməsi mərhələsi daxil edilmir.



Şəkil 7. Five A's metodologiyasının mərhələləri

Marba'n metodologiyası

Bu metodologiya (Marba'n et al., 2008) Data Mining prosesini tədqiqat prosesi və mühəndislik problemlərinin həlli baxımından nəzərdən keçirmək fikrinə əsaslanır. Buna görə Data Mining sistemlərinin inkişaf prosesləri özündə CRISP-DM standartına daxil olmayan lakin mühəndislik fəaliyyətində zəruri olan tapşırıq və hərəkətləri cəmləşdirir. Bu məqsədlər üçün proqram təminatı sahəsində iki təsdiq edilmiş standart əsaslanmaq təklif olunur: IEEE 1074 (IEEE, 1991) və ISO 12207 (ISO, 1995). Bura təlim və inkişaf prosesləri daxildir. İnkişaf prosesləri aşağıdakılara əsaslanır:

- məlumatların öncədən emalı;
- məlumatların emalı, əsas KDD&DM olan proses (seçmə, götürmə, çevrilmə, Data Mining, şərh / qiymətləndirmə);

- emaldan sonrakı nəticələr (burada, proqram mühəndisliyində olan ənənəvi addımlarından istifadə edilir).

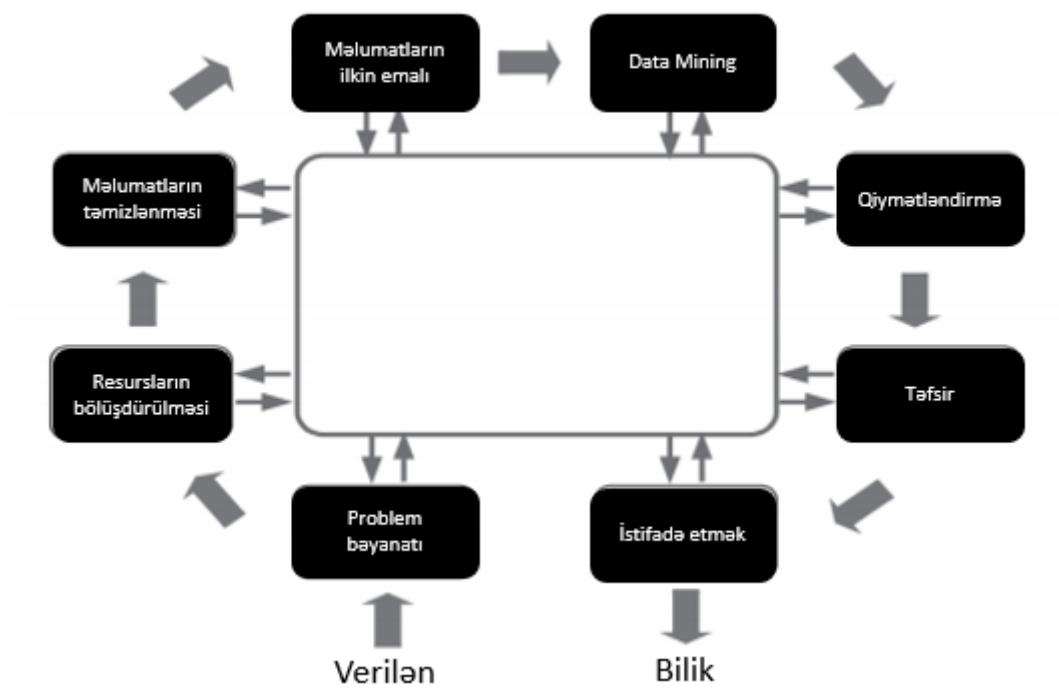
KDD Roadmap metodologiyası

KDD Roadmap (Debusse et al., 2001.) Witness Miner toolkit tərəfindən istifadə olunan məlumatların intellektual təhlili metodologiyasıdır.

Şəkil 8.-də göstərildiyi kimi, KDD Roadmap təkrarlanan metodologiyadır və səkkiz addımdan ibarətdir:

1. problem bəyanatı;
2. resursların bölüşdürülməsi;
3. məlumatların tənzimlənməsi;
4. məlumatların ilkin emalı;
5. Data Mining;
6. qiymətləndirmə;
7. təfsir;
8. istifadə etmək;

KDD Roadmap-ın əsas töfhəsi biliklərin aşkar edilməsi prosesinin səmərəliliyini artırmaq üçün tapşırıqların nəzərdən keçirilməsinə ehtiyat mənbəyinin daxil edilməsidir.



Şəkil 8. KDD Roadmap metodologiyasının mərhələləri

KDD & DM prosesinin standartları davamlı şəkildə inkişaf edir və müasir standartları tamamlayan yeniləri meydana gəlir. Bu metodologiya məlumatların intellektual təhlili sistemlərinin tədqiqat üçün onlardan istifadə etməyə imkan verən KDD&DM metodologiyalarının kifayət qədər yetkinliyini göstərir.

2.2. İqtisadi məlumatların intellektual təhlil üsullarının icmalı

Məlumatların intellektual təhlili sistemləri gizli, əvvəllər bilinməyən, mənalı və potensial faydalı nümunələri böyük həcmdə heterogen, mürəkkəb quruluşlu məlumatların axtarışını avtomatlaşdırmaqdan ibarət olan proqram qərarlarına dəstək sistemlərinin bir sinfidir.

Data Mining sistemlərinin inkişafının əsasları

Məlumatların intellektual təhlili sistemlərinin və Data Mining tapşırıqlarını həyata keçirən sistemlərin yaradılmasının əsas məqsədi, mövzu sahəsinin obyektlərinin vəziyyəti və inkişaf meyilləri haqqında yeni biliklər əldə etmək üçün mürəkkəb, bir qədər də rəsmiləşdirilmiş strukturlar tərəfindən formalaşan məlumat növlərinin işlənməsi və idarə olunmasıdır.

Data Mining əsasında avtomatlaşdırılmış, məlumat hasilatına yönəlmiş sistemlərdə aşağıdakı əsas modullar fərqləndirilir: məlumatların işlənməsi modulu, identifikasiya modulu, tənzimləmə və qiymətləndirmə modulu, təhlil və nəticələr əldə etmək üçün Data Mining metodik tətbiqi modulu, mənbə məlumatları və əldə edilmiş nəticələr olan anbar, sistem biliklərinin mütəşəkkil dəstini və problem sahəsini təmsil edən məlumat bazası.

Data mining sistemlərinin tədqiqatçısı ilk növbədə aşağıdakı bilik təbəqələri ilə ifadə olunan sistemin semantikasını müəyyənləşdirməlidir: [17]

1. Bilik haradadır. Sistemin əhatə dairəsi, məkan məhdudiyyətləri, şərtlər, digər sistemlərlə əlaqə.
2. Bilik kimdir. Sistem istifadəçiləri, səlahiyyətçiləri və istifadəçi tələbləri.
3. Niyə bilik. Sistemin tətbiq məqsədi və vəzifələrinin cəmi.
4. Bilik nədir. Giriş və çıxış sistemi məlumatları. Onların növləri, quruluşları, modelləri.
5. Bilik kimi. Sistemin fəaliyyətini təyin edən funksiyalar, əməliyyatlar, alqoritmlər.
6. Nə üçün bilik. Sistemin əhəmiyyəti, aktuallığı.
7. Nə vaxt bilik. Vaxt məhdudiyyəti.

8. Nə qədər bilik. Kəmiyyət göstəriciləri, xərclər, fəaliyyət göstəriciləri.

9. Hansı bilik. Sistemin keyfiyyətə qiymətləndirilməsi.

Qeyd edək ki, bilik təbəqələri sistemin mümkün məhdudiyyətlərlə giriş-çıxış (sistem 1, 2, 4, 7), funksiyalar (təbəqə 3, 5), meyarlar (təbəqə 6, 8, 9) kimi obyektlərinə birləşdirilə bilər. Bu genişlənmiş bilik təbəqələri konseptual tədqiqat üçün əsas götürülə bilər.

Data Mining S sisteminin konseptual tədqiqat çərçivəsində dil trixotomiyasına yönəlmiş “İşarə-Məna-İşarə etmə” adlı modelin tərifini nəzərdən keçirək.

S sisteminin ümumiləşdirilmiş komponent modeli, o cümlədən giriş X , çıxış Y , girişi F çıxışına çevirmək üçün bir çox model, F modelinin keyfiyyətini təyin edən parametrləri qiymətləndirmək üçün istifadə olunan bir çox meyar, $Cs = \langle Y, X, F, K \rangle$ kimi göstərilə bilər. Bu ənənəvi sistem modeli müəyyən bir mövzu sahəsindəki vəzifələr üçün ən azı iki səviyyəli bir nümayəndəliyə sahib ola bilər: bir səviyyədə - son istifadəçi baxımından, digər səviyyədə - riyazi modellər və metodlar baxımından.

Cs modelinin hər bir komponenti daxili və xarici şərtləri olan semantika ilə xarakterizə olunur. Semantikanın xarici təfsiri modelin S intensivliyini son istifadəçi baxımından ifadə olunan hər bir komponentin linqvistik nümayəndəliyi şəklində müəyyən edir. Daxili isə model S genişlənməsini hər bir komponent üçün etibarlı əməliyyatları yerinə yetirmək qaydaları, münasibətlərin cəmi və model komponentləri arasında riyazi asılılıq kimi müəyyənləşdirir. Beləliklə, daxili semantika analiz problemlərinin həlli üçün modellər tərəfindən müəyyən edilir və "Dəyərlər" yaratmaq üçün istifadə olunur. Xarici semantika isə mövzu sahəsinin mövzu-obyekt xəritəsini və S modelinin komponentlərində formal sistemlərin tədqiqatçısına məlum olan obyektləri və eyni zamanda “İşarə” ni müəyyənləşdirir.

S sisteminin model Cs üçün daxili semantikasını təyin edən $Rs = R(Y, X, F, K)$ və xarici semantikasını təmsil edən $Ps = P(Y, X, F, K)$ funksiyalarını təyin edək.

Onda mücərrəd Data mining S sisteminin konseptual modeli $Ms = \langle Cs, Rs, Ps \rangle$ kimi müəyyən edilə bilər və bu model “İşarə-Məna-İşarə etmə” dil trixotomiyasını təqdim edir. Məlumatların müəyyən bir sinfini təhlil edərkən S

sisteminin komponentlərinin və semantik modellərinin müqayisəsi, ümumiyyətlə ətraf mühitin və onların təcrübələrinə əsaslanaraq ən uyğun qərarı verə bilən sistem tədqiqatçısı tərəfindən həyata keçirilir.

Müəssisələrin iqtisadi səmərəliliyinin ekspress təhlili sistemi

Zaman sırası proqnozu və qeyri-səlis meyllərin təhlili əsasında müəssisələrin ekspress təhlili üçün internet xidməti elektron xidmətlər bazarı segmentinə yönəldilmişdir. Müəssisələrin sürətli təhlili üçün texniki və iqtisadi göstəricilərin modelləşdirilməsi və proqnozlaşdırılması neyron şəbəkələrinə əsaslanan bir neçə qeyri-səlis modeli özündə birləşdirən inteqral metod əsasında həyata keçirilir. [1]

İnternet xidmətində müəssisələrin ekspress təhlilinin əsas vəzifələri aşağıdakılardır:

- Açıq mühasibat hesabatlarına əsasən texniki və iqtisadi göstəricilərin hesablanması və proqnozlaşdırılması.

- Texniki və iqtisadi göstəricilər baxımından inkişaf meyllərinin modelləşdirilməsi, proqnozlaşdırılması və lingvistik ümumiləşdirilməsi.

- İqtisadi təhlil və ekspress analiz nəticələrinin təfsiri.

İnternet xidmətində tətbiq olunan müəssisənin ekspress təhlili metodu müəssisənin iqtisadi təhlilində ən çox istifadə olunan dörd qrup göstəricinin hesablanmasını təmin edir:

1. satılma və ödəmə qabiliyyətinin göstəriciləri;
2. maliyyə müstəqilliyinin göstəriciləri;
3. gəlirlilik göstəriciləri;
4. işgüzar fəaliyyətin göstəriciləri.

İstifadəçinin davranışını izləmək üçün Data Mining sistemi

Burada ki, vəzifə müəyyən bir sistem istifadəçilərinin özlərinə qarşı ola biləcək təhlükələri müəyyənləşdirmək üçün onların davranışlarını izləyən bir sistem hazırlamaqdır. Tədqiqatların nəticələrinə görə ilkin fərziyyə ondan ibarətdir ki,

istifadəçilər və proqramların fəaliyyəti tamamilə izlənilə bilər eyni zamanda təhdidləri təsnif etmək üçün uyğun model qurmaq lazımdır.

Müdaxilənin aşkarlanması tapşırıqlarını həll edən sistemin xüsusiyyətləri aşağıdakılardır:

1. tarixi məlumatların toplanması;
2. normal davranış və ya müdaxiləni modelləşdirmək;
3. istifadəçi davranışında anomaliyaları axtarmaq;
4. qurulmuş modellərə görə sistemin cari fəaliyyətini təsvir edən hadisələrin təsnifatı və təhlili.

Texnoloji prosesin keyfiyyətinin təhlili və proqnozlaşdırılması

İstehsal olunan məhsulların keyfiyyətini təhlil edərkən və proqnozlaşdırarkən, istehsal prosesinin hansı parametrlərinin məhsulun keyfiyyətinə təsir etməsi barədə sual yaranır. Bu problemi həll etmək üçün qeyri-səlis dəyişənlərə, təkamül metodlarına və qeyri-səlis qərar ağaclarına əsaslanan alqoritm ilə aşağıdakı şərtləri ödəyən sistem hazırlanmışdır: [10]

1. "əgər ... sonra ... əksinə" qeyri-səlis qaydalar sistemi şəklində təqdim olunan məhsul keyfiyyətinin istehsal prosesinin xüsusiyyətlərindən asılılığı modelini qurur;
2. istehsal prosesinin xüsusiyyətlərinə görə məhsulun keyfiyyət göstəricilərini proqnozlaşdırır;
3. prosesin xüsusiyyətlərini keyfiyyətə təsir dərəcəsinə görə sıralayır.

Data Mining sisteminin tədris prosesinə tətbiqi

İki sinif obyektinə nəzərdən keçirilir: O1 tələbələr toplusudur, O2 isə fənlər toplusudur. Hər bir tələbə O1_j iki növ atribut dəsti ilə xarakterizə olunur. Birinci növ, fənlərdəki ballar ilə əlaqələndirilir. Atributun ikinci növü intizamın öyrənilməsi prosesində onun təhsil mənbələri ilə qarşılıqlı əlaqəsinin faizini təsvir edir. Hər O2 fənni aşağıdakı atributlarla təsvir olunur:

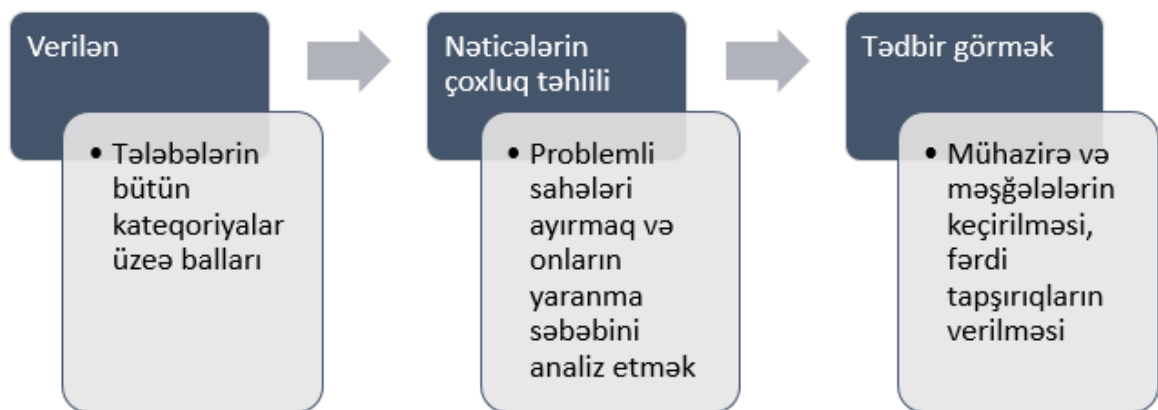
kəmiyyət "5", kəmiyyət "4", kəmiyyət "3" və kəmiyyət "2".

Problemin quruluşuna iki xüsusi tapşırığın həlli daxildir:

Birinci tapşırığın məqsədi, təhsil texnologiyalarının daha da tənzimlənməsi üçün uğurlu, sabit və ya uğursuz tələbələrin atributlara görə qruplara bölünməsinə tapmaqdır.

İkinci tapşırığın məqsədi, fənlər qruplarını tələbələr tərəfindən inkişaflarının orta xüsusiyyətləri ilə fərqləndirmək kimi tərtib edilmişdir.

Şəkil 9.-da klasterləmə metodunu reallaşdıran alt sistemin modeli göstərilir.



Bu sistemin tətbiqi tələbə fəaliyyəti haqqında məlumatları rahatlıqla saxlamağa, qrafik bir təqdimatı tez əldə etməyə imkan verir. Eyni zamanda intellektual analizin nəticələrinə əsaslanaraq tədris prosesini yaxşılaşdırmağa yönəlmiş sistemdir.

Data Mining-in kompüter oyunlarına tətbiqi

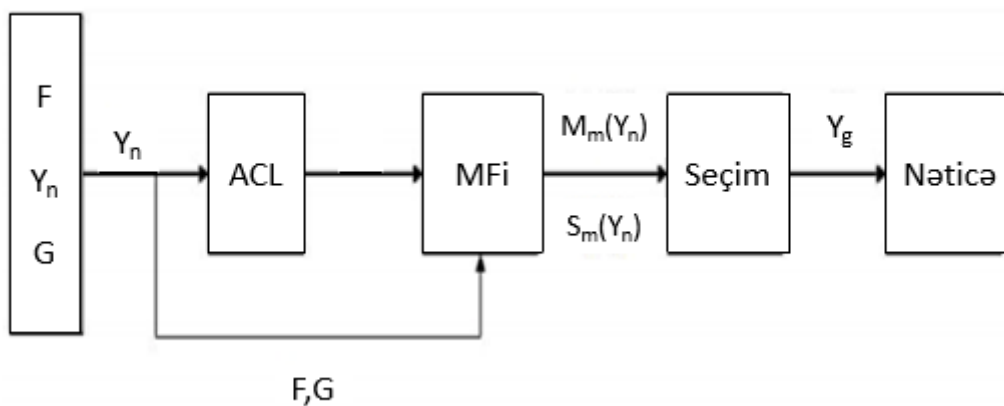
Əksər kompüter və mobil oyunlarda, hər oyun seansının sonunda oyunçuya nəticələri haqqında məlumat verilir. Kompüter oyunlarında reytinglərdən əlavə olaraq, fərdi rəy əsasında oyunçuya daha dolğun məlumat vermək vəd edilir.

İnkişaf etmiş sistemin tətbiqi oyun sinifinin istifadə etdiyi oyunçunun effektivliyini qiymətləndirərək oyun seanslarının səmərəliliyini artıracaqdır. [13]

Zaman sırası ilə proseslərin proqnozlaşdırılması sistemi

Zaman sırasının proqnozu fərdi prosesin proqnozunu əldə etmək üçün hazırlanmışdır. Bu cür problemlərin həlli iqtisadiyyatda, elektrik enerjisində, tibbdə, texnologiyada və idarəetmədə tələb olunur.

Şəkil 10. intellektual analiz metodlarından istifadə edərək zaman sıralarını proqnozlaşdırmaq üçün hazırlanmış sistemin mücərrəd modelini göstərir. Bu tətbiqin mərkəzində olan zaman sırası analizinə yanaşmanın aktuallığını, ən təsirli proqnoz əldə etmək üçün klassik statistik analiz metodlarını və qeyri-səlis zaman sırası nəzəriyyəsini birləşdirmək durur. [3]



Şəkil 10. Zaman sırası proqnozlaşdırılması üçün xidmət məlumat axını cədvəli (burada Y_n ədədi, F - hamarlaşdırma funksiyası, G – üfüqi proqnoz, $M_m(Y_n)$ – zaman sırasının proqnozlaşdırılması nəticəsində əldə edilən zaman sırası modelləri, $S_m(Y_n)$ – zaman sırası qalıqlarının proqnozlaşdırılması nəticəsində əldə edilən zaman sırası qalıqlarının modelləri, Y_g – seçilmiş zaman sırası proqnozu)

Zaman sırası şəklində məlumatların emalı mərhələləri aşağıdakı blok modullarından ibarətdir:

1. ACL miqyaslı modul. Mənbə məlumatlarını qeyri-səlis modellərə çevrilib öncədən emal etmək üçün hazırlanmışdır.

2. Riyazi proqnoz modellərinin modulu (MFi). Burada müxtəlif elmi yanaşmaların zaman sırası modelləri saxlanılır və qiymətləndirilir: qeyri-səlis modellər və statistik modellər.

3. Seçim modulu. Müxtəlif meyarlardan istifadə edərək proqnozlaşdırma nəticələrinə əsaslanan ən yaxşı modeli seçmək modulu.

4. Nəticə modulu. Görüntülmə və nəticələrin xülasəsi modulu.

Birinci mərhələdə zaman sıralarının qeyri-səlis həyata keçirilməsi baş verir.

İkinci mərhələdə qeyri-səlis trend modelləri qurulur və proqnozlaşdırılan dəyərlər paralel proqnoz üfünə uyğun olaraq hesablanır. Bu ARIMA modelindən istifadə edərək blokdakı zaman sıralarının qalıqlarını proqnozlaşdırmaq üçün istifadə olunur. Blokun çıxışında əsas zaman sıralarının proqnoz modelləri, eləcə də bir çox qalıq modelləri meydana gəlir.

Üçüncü mərhələdə ən yaxşı proqnozlaşdırma modelinin seçimi zaman sırası və qalıqlarının əldə edilmiş proqnozlarının keyfiyyət meyarlarının təhlili əsasında aparılır.

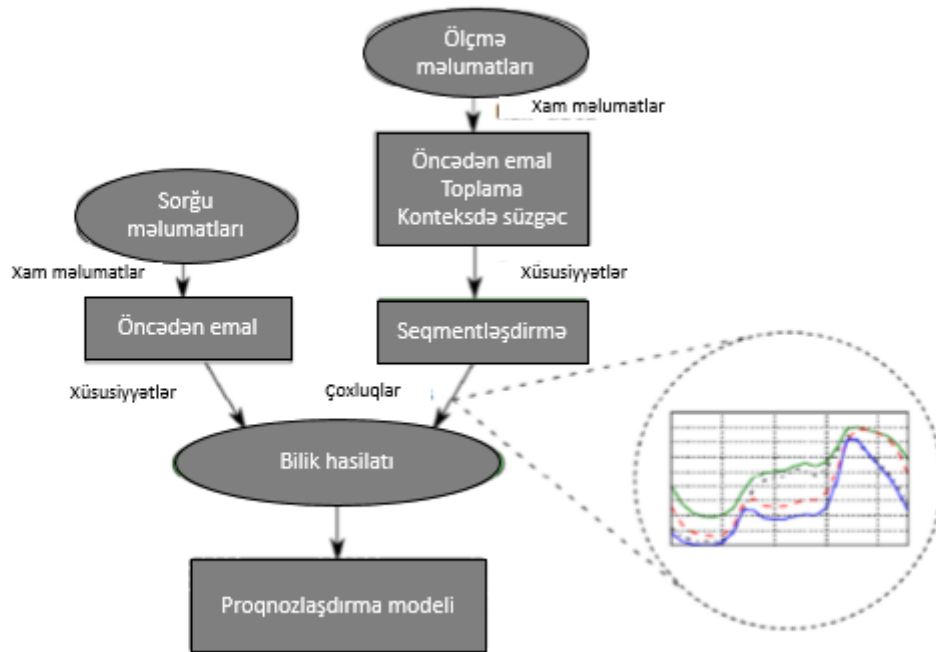
Bazarda yeni bir obyektin dəyərini qiymətləndirən sistem

Bazara girən bir obyektin qiymətinin hesablanması aktual problemdir. Obyektlərin dəyərinin parametrləri toplusundan asılılığının qurulması vəzifələrində, məlumat hasilatını həyata keçirən bir tətbiqə baxaq. Asılılığı qurmaq üçün lazım olan ilkin məlumatlara parametrləri və məlum dəyəri olan obyekt toplusu daxildir. Yeni obyektin dəyərini müəyyənləşdirmək problemi asılılığın qurulmasının iki üsulu ilə həll edildi: xətti çoxsaylı reqressiya əsasında və xətanın geri yayılması metodu ilə öyrədilən neyron şəbəkəsindən istifadə etməklə. [4]

Öyrənilmiş neyron şəbəkəsi bir faylda saxlanılır və bundan sonra yeni bir obyektin dəyərini hesablamaq üçün ondan istifadə olunur.

Enerji istehlakının proqnozlaşdırılması

Son istehlakçıların elektrik enerjisi istehlakına dair toplanmış məlumatlar bizə enerji təchizatı şirkətləri üçün faydalı olan proqnozlar verməyə imkan verir.



Şəkil 11. Enerji istehlakının proqnozlaşdırılması arxitekturası

İlkin məlumatlar iki yerə bölünür:

- istehlakçı ailələrinin anket məlumatları (sorgu məlumatları). Buraya eyni mənzildə yaşayan istehlakçılar (yaş, miqdar, ailə gəlirləri) və istifadə olunan məişət cihazlarının siyahısı (soyuducu, kompüter, kondisioner və s.) daxildir;
- enerji ölçmə cihazlarının oxunuşları (ölçmə məlumatları).

Sistemdə aşağıdakı məlumatların işlənməsi üsulları tətbiq olunur:

1. Əks dəyərləri bərpa etmək, istehlakçı qruplarının xüsusiyyətlərini (gəlir baxımından, ailə üzvlərinin sayı və s.) seçmək üçün məlumatların əvvəlcədən işlənməsi.
2. Mövsüm (məsələn, yay və ya qış) oxunuşlarını, gündüz / axşam oxunuşlarını seçmək üçün məlumatların süzülməsi (kontekstdə süzülmə).
3. 24 saatlıq trafik enerjisi istehlakının formalaşması (yığılması).
4. Çoxluq əldə etmək üçün 24 saatlıq elektrik istehlakı trafikinin seqmentləşdirilməsi. Bir seqmentləşdirmə alqoritmi olaraq fuzzy c-means klasterləşmə metodundan istifadə olunur.

5. İstehlak obyektlərinin istehlakçı ailələrinin xüsusiyyətlərindən və enerji istehlakı qruplarından asılılıq modellərinin çıxarılması.

6. Çıxarılan məlumatlara əsasən alınan hər bir klasterdə elektrik istehlakı nümunələrinin proqnozlaşdırılması.

Proqnozlaşdırmaq üçün üç fərqli metoddan istifadə edilir: reqressiya, qeyri-səlis Takagi-Sugeno proqnozlaşdırma modeli və dəstək vektoru metoduna əsaslanan təsnifat. [27]

Data Mining tapşırıqlarına giriş

Ümumiyyətlə, hər hansı bir problemin ifadəsi ən ümumi giriş-çıxış sistem modeli kimi təqdim edilə bilər:

$$Y=F(X) \quad (1)$$

Burada, Y istənilən nəticə, X mənbə məlumatlarının məcmusu (verilənlər bazası), F mənbə məlumatlarını istədiyiniz nəticəyə çevirən sistemdir.

Adətən, ilkin məlumatlar və tələb olunan nəticə domen mütəxəssisi tərəfindən təyin edilir və Data Mining sisteminin tədqiqatçısının vəzifəsi onu tərtib etmək və həyata keçirməkdir. Bunun üçün tədqiqatçı Data Mining tapşırıqlarının müvafiq sinifini təyin etməli, ilkin məlumatları təmsil etmək üçün modelləri, məlumatların işlənməsi və saxlanması metodlarını seçməli və F sisteminin arxitekturasını əsaslandırmalıdır. Qeyd edək ki, Data Mining probleminin həlli üçün, metod (alqoritm) seçimi üçün mənbə və çıxış məlumatlarının təqdimatı, dəqiqliyi və vaxtı həlledicidir. Tez-tez bir neçə rəqabət üsulu seçilir və aralarında ən effektiv (dəqiqlik / vaxt meyarına görə) olanın axtarışı aparılır.

Bir çox obyekt özünün atributlar dəsti kimi saxlayan verilənlər bazasından məlumat çıxarmaq üçün istifadə olunan Data Mining-in əsas vəzifələrinin məqsədini və formal tərtibini nəzərdən keçirək.

(1) ifadəsindəki ilkin verilənlər bazası aşağıdakıları ehtiva edə bilər:

1. Xam məlumatlar (raw data). Müşahidə və ya ölçmə məlumatlarıdır.
2. Bəzi alqoritmlər nəticəsində əldə edilən çevrilmiş məlumatlar.

3. Xüsusiyyətlər (feature). Bunlar məlumatların bəzi tərəflərini xarakterizə edən xam məlumatlara çevrilir.

4. Xam məlumatların xüsusiyyətlərini izah edən modellərin parametrləri.

5. Problemi həll etmək üçün əlavə məlumatlar.

Data Mining problemlərinin formal qoyuluşunu nəzərdən keçirərkən, məlumatların nümayişi variantlarını nəzərə almadan C. C. Aggarwal tərəfindən təklif olunan verilənlər bazasının iki ölçülü nümayəndəliyinə yönəlmiş yanaşmadan istifadə edirik.

Tərif 1.1. MD ($n \times d$) çoxölçülü verilənlər bazası n qeydlərin $X_1, \dots, X_n$ toplusudur, harada ki, hər bir X_i ($i = 1, 2, \dots, n$) qeyd, d isə atributların sayı olduğu halda $(x^1_i, \dots, x^d_i)$ qeydindən ibarətdir.

Münasibətlər nəzəriyyəsindən istifadə edərək $Y=F(X)$ ifadəsi üçün mənbəli məlumat bazasının tərifini xüsusi kontekst şəklində verə bilərik.

Tərif 1.2. Çoxqiymətli kontekst dedikdə aşağıdakı dördlü nəzərdə tutulur:

$$K=(G, M, W, J)$$

Burada, G obyektlər toplusudur (MD verilənlər bazasının sətirləri), M çoxqiymətli atributlar toplusudur (MD verilənlər bazasının sütunları), W atribut dəyərlərinin məcmusudur və J elə bir münasibətdir ki, $I \subseteq G \times M \times W$, hər hansı bir $g \in G$, $m \in M$ üçün $(g, m, w) \in J$ – ni ödəyən ən azı bir $w \in W$ var. $(g, m, w) \in J$ ifadəsi g obyektinin w mənalı m atributunun olduğunu göstərir. [24]

Aşağıdakı hissələrdə Data Mining-in əsas vəzifələrinin formalaşdırılmasına baxa bilərik.

Klasterləşmə və seqmentləşmə vəzifələrinin qoyuluşu

Klasterləşdirmə və seqmentləşdirmə, müəyyən bir metrikaya yaxın dəyərlərə malik olan uyğun, oxşar obyektlərin axtarışı və qruplaşdırılması üçün nəzərdə tutulmuşdur.

Seqmentasiya ilkin məlumatlar həmçins və bir vektorla təmsil olunduqda tətbiq olunur. Daha mürəkkəb bir vəziyyətdə, yəni məlumatlar həmçins və çox ölçülü bir formada təqdim edildikdə çoxluq təhlili istifadə olunur.

Tərif 2.1. MD($n \times d$) verilənlər bazası üçün sətirlər üzrə $C_1, \dots, C_k$ kimi çoxlu qruplara bölünməsinə elə müəyyənləşdirmək lazımdır ki, hər bir oxşar qrupda oxşar sətirlər, fərqli qruplarda isə fərqlilər olsun.

$X = (x_1, \dots, x_d)$ və $Y = (y_1, \dots, y_d)$ kimi iki çoxluğu nəzərdən keçirək. Hər bir atribut üçün hesablanmış oxşarlıq ölçüləri $S(x_i, y_i)$ cəminə bərabər olduğundan X və Y arasındakı oxşarlığı hesablamaq üçün bir üsul təklif etmək təbiidir.

Ölçməni seçmək üçün ən sadə iki yol var. Əgər $x_i = y_i$ olarsa $S(x_i, y_i) = 1$, əks halda $S(x_i, y_i) = 0$ olacaq. Ancaq nadir hallarda görünən oxşarlıqlar və ya fərqlər statistik baxımdan daha əhəmiyyətlidir. Kateqoriyalı məlumatlar kontekstində oxşarlıqların hesablanması üçün statistik məlumat xassələrinin istifadəsi vacibdir. Belə qlobal statistik məlumat xassələri, oxşarlığı daha dəqiq ölçmək üçün Mahalanobis metriyasında tətbiq olunur. Bu metrikanın fikri odur ki, kateqoriyalı atributların qeyri-standart dəyərləri ümumi olan dəyərlərdən daha böyük əhəmiyyətə malik olsun.

Təsnifat vəzifələrinin qoyuluşu

Təsnifatın məqsədi əvvəlcədən təyin olunmuş siniflərə (qruplara) görə bir çox obyektə elə paylamaqdır ki, hər sinifdə oxşar obyektlər olsun. Klasterləşmədən fərqli olaraq, təsnifat problemi təlim dəstinin olması şərtilə formalaşdırılır. Bu isə obyektlərin məlumat bazası MD($n \times d$) və onun hər bir obyektə üçün müəyyən dəstdən sinif etiketinin olması ilə müəyyən edilmişdir.

Təsnifat, tanınmış nümunələrdən istifadə edərək müəllimlə təlim tələb edən vəzifələrə aiddir. Müəllimlə öyrənən zaman, MD ($n \times d$) verilənlər bazası və onunla əlaqəli sinif etiket vektoru $Cl(n)$ iki dəstə bölünür:

- Model (təsnifat) hazırlamaq üçün istifadə edilən təlim dəsti, verilmiş siniflərlə nümunələr toplusu.
- Test dəsti, bu dəst nümunələrin giriş və çıxış dəyərlərini ehtiva edir. Burada çıxış dəyərləri (sinif etiketləri) təsnifatçı keyfiyyətini yoxlamaq üçün istifadə olunur.

Təlim və test dəstlərindəki təsnifatın düzgünlüyü (və səhvləri), alınan sinif etiketlərinin məlum olanlarla uyğunluğu (və təsadüf olmaması) meyarı ilə

qiymətləndirilir. Bu məqsədlər üçün hətta kross-yoxlama (çarpaz yoxlama) da tətbiq olunur:

a) test dəstinin təsnifat dəqiqliyi təlim dəstinin təsnifat dəqiqliyi ilə müqayisə edilir.

b) test dəstinin təsnifatı təlim dəstinin təsnifatı ilə təxminən eyni dəqiqlik nəticəsi verərsə, bu təsnifat modelinin çarpaz şəkildə yoxlanıldığı hesab olunur.

Təsnifatçının düzgünlüyünü qiymətləndirmək üçün onu təlim və test dəstlərinə ayırmaq tövsiyə olunur, ardınca isə müəyyən nisbətdə orta təsadüfi seçmə aparıla bilər. Məsələn, təlim dəsti məlumatların üçdə ikisini, test məlumat dəsti isə məlumatların üçdə birini ehtiva edir.

Tərif 3.1. İlk verilənlər bazası $MD (n \times d)$, sinif nişanlarının əlaqəli vektoru $Cl(n)$ və n xətlərinin hər biri ilə əlaqələndirilən (MD-yə giriş) etiketi müəyyənləşdirmək (proqnozlaşdırmaq) üçün $(n+1)$ ölçülü $Cl(n+1)$ etiketli bir model(təsnifatçı) MA yaratmaq lazımdır.

Təsnifat sinif etiketi olaraq göstərilən bir dəyəri təxmin etmək üçün də istifadə edilə bilər.

Proqnozlaşdırma vəzifələrinin qoyuluşu

Təsnifat probleminin yuxarıda göstərilən formulasiyasından fərqli olaraq, proqnozlaşdırma problemi daha çox ümumi formada, məlumatların ya zaman sabitliyindən, ya da əvvəlki dəyərlərdən asılılığını müəyyənləşdirmək vəzifəsi şəklində tərtib edilmişdir. Bu vəziyyətdə verilənlər bazası zamana görə, sifariş verilən dəyərlər ardıcılığı şəklində nəzərdən keçirilir və təlim nümunələri kimi istifadə olunur. Proqnozlaşdırma probleminin həllinin nəticəsi müntəzəm dövrlər üçün verilənlər bazasında yeni qeydlərdir.

Tərif 4.1. Verilənlər bazası $MD (n \times d)$ və zaman anlarının $Y(n)$ ilə əlaqəli vektoru üçün $Y(n+1)$, $Y(n+2)$, ..., $Y(n)$ dəyərləri üçün MF modeli (alqoritmi) yaratmaq tələb olunur, hansıki $Y(n+1)$, $Y(n+2)$, ..., $Y(n+p)$ dəyərləri üçün $MD((n+1) \times d)$, $MD((n+2) \times d)$, ..., $MD((n+p) \times d)$ matrislərinin sətirlərinin dəyərini hesablaya bilər. P dəyişəni proqnoz üföqü adlanır. Bu vəziyyətdə, matrisə

MD ($n \times d$) $Y(n)$ vektoru ilə birlikdə MF modelinin qurulması üçün təlim dəsti təşkil edir.

Proqnozlaşdırma modelinin düzgünlüyünü qiymətləndirmək üçün alınan dəyərin həqiqi modeldən kənarlaşması tədbirləri istifadə olunur. Təsnifat probleminə olduğu kimi, proqnozlaşdırma probleminə də k dəyərli çarpaz təsdiqləmə tətbiq edilə bilər.

Beləliklə, proqnozlaşdırma bəzi asılılıq modeli əsasında tarixən toplanmış məlumatlardan yeni, əvvəllər məlum olmayan bir dəyəri hesablamaq problemini həll edir. Bu model parametrik funksiya təqdim edə bilər, bu parametrlər isə qiymətləndirilməli və ya məlumatlardan alınan ardıcılığın bəzi qaydalarına əsaslanmalıdır.

Assosiativ qaydaların axtarışı vəzifələrinin qoyuluşu

Assosiativ qaydaların axtarışı MD ($n \times d$) verilənlər bazası sətirlərində tez-tez və birgə meydana gələn qiymətlər çoxluğunu tapmağa yönəlmişdir. Tez-tez meydana gələn qiymətlər çoxluğunun tərifini təqdim edək.

Tərif 5.1. İkili matrisa MD($n \times d$) üçün bütün sütun qiymətləri 1-ə bərabər olan altçoxluqlar təyin etmək tələb olunur və bu da matrisin sətirlərinin s -ci hissəsinin qiymətindən az deyil. Bu nümunədə s -in mənası minimum dəstək adlanır.

Verilənlər bazasında assosiasianın tapılması vəzifəsi, MD($n \times d$) matrisinin sətirlərində tez-tez meydana gələn assosiativ qaydaları çıxarmaq şəklində tərtib edilmişdir.

Minimum dəstək tələblərinə cavab verən qiymətlər çoxluğu (və ya nümunələri) bir-birinə yaxın adlanır. Bir-birinə yaxın çoxluqlar etibarlılıq olaraq bilinən bir ölçü istifadə edərək birlik qaydaları yaratmaq üçün istifadə edilə bilər.

Tərif 5.2. $X \Rightarrow Y$ qaydasının etibarlılığı MD($n \times d$) matrisasının sətrində X çoxluğunun qiymətlərinin də daxil olması şərti ilə Y çoxluğunu ehtiva edən şərti ehtimaldır. Bu ehtimal, $X \cup Y$ birləşməsinin X çoxluğuna bölməsinin nəticəsinə bərabər olan kəmiyyət ilə qiymətləndirilir.

Çıxarılan qaydanın faydalılığını qiymətləndirmək üçün yaxşılaşdırma adlı bir göstərici hesablanır. Bu təsadüfi təxmin qaydasının daha faydalı olub olmadığını mühakimə etməyə imkan verir.

Tərif 5.3. $X \Rightarrow Y$ qaydasının yaxşılaşdırılması, $X \cup Y$ birləşməsinə ehtiva edən $MD(n \times d)$ matrisinin X çoxluğunun sətirləri sayının Y çoxluğunun sətirləri sayına nisbətidir.

Qaydanın təkmilləşdirilməsi birdən çoxdursa, bu o deməkdir ki, Y çoxluğunun olacağını proqnozlaşdırmaq üçün $X \Rightarrow Y$ assosiativ qaydadan istifadə təsadüfi təxminlərdən daha çoxdur, birdən az olduqda isə əksinə.

Tərif 5.4. Tutaq ki, A və B ikilisi $MD(n \times d)$ matrisinin sətirlərində iki çoxluq olsun. Onda qayda $A \Rightarrow B$ aşağıdakı şərtlər ödəndikdə s dəstək səviyyəsində və c güvən səviyyəsində etibarlıdır.

1. A çoxluğunun dəstək səviyyəsi s -dən az deyil.
2. $A \Rightarrow B$ qaydalarına güvən səviyyəsi c -dən az deyil.
3. $A \Rightarrow B$ qaydalarının təkmilləşdirilməsi 1-dən çoxdur.

Birlik qaydaları müəyyən dərəcədə "əgər-onda" tipli bir model kimi qəbul edilə bilər. Bu vəziyyətdə, çıxarılmış assosiativ qaydalardan A çoxluğuna görə B çoxluğunu proqnozlaşdırmağa imkan verən predikativ bir model kimi istifadə etmək faydalıdır.

Anomaliyaların axtarışı və aşkarlanması vəzifələrinin qoyuluşu

Anomaliyaları (və ya istisnaları) tapmaq və aşkar etmək vəzifəsini nəzərdən keçirək. Ümumiyyətlə, bu vəzifə aşağıdakı variantlarda tərtib edilə bilər:

1. Müşahidə olunan məlumatlar kontekstində yaranan anomaliyaları tapmaq və aşkar etmək vəzifəsi. Bu vəzifə məlumatların məhdudiyyətlər, icazə verilən dəyərlər, xassələr və ya davranışlar barədə məlumatı olması şərtlə tərtib edilmişdir.

2. Mövcud anomaliyaları olan və anomaliyaları olmayan məlumatlarla əlaqəli, məlumatların mövcudluğu şərti ilə anomaliyaları axtarmaq və aşkar etmək vəzifəsi (məlumatların təsnifatı tapşırığı).

3. Qeyri-müəyyənlik qarşısında anomaliyaları tapmaq və aşkar etmək vəzifəsi.

Anomaliyaları axtarmaq və aşkar etmək probleminin son variantı, anomal cismin nadir və tipik olmayan xüsusiyyətlərinin olması ilə xarakterizə olunduğu ehtimalına əsasən formalaşdırıla bilər.

Anomaliyalar tək (xarici) və çox kimi təsnif edilir. Aşağıda qeyri-müəyyənlik altında vahid xarici tipli anomaliyaların axtarışı və aşkarlanması üçün problem bəyanatının tərtibatı verilmişdir.

Tərif 6.1. MD ($n \times d$) verilənlər bazası üçün sətirləri qalan sətirlərdəki xüsusiyyətlərdən əhəmiyyətli dərəcədə fərqlənəcək şəkildə müəyyənləşdirmək lazımdır. Bu cür xüsusiyyətlərə isə anomaliyalar deyilir.

Bu tərifdən göründüyü kimi, anomaliyaların axtarışı sətirlər arasında əhəmiyyətli uyğunsuzluqların aşkarlanması baxımından və məlumatların yığılmasına əsaslanaraq nəzərdən keçirilə bilər. Digər tərəfdən də xüsusiyyətlərin nadirliyini axtarmaq vəzifəsi assosiativ qaydaların axtarışı nəticəsində nəzərdən keçirilən, bir-birinə yaxın nümunələrin axtarılması vəzifəsinə tərsdir.

Rəsmi konseptual analiz əsasında konsepsiyaların axtarışı vəzifələrinin qoyuluşu

Formal konsepsiyaların təhlili (Formal Concept Analysis (FCA)) anlayışlar arasında əlaqə qurmaq üçün hazırlanmışdır. Daha konkret bir vəziyyətdə, ümumi xüsusiyyətlərə malik olan, lakin oxşar dəyərlərə malik olmayan məlumat qruplaşmaları yaratmaq lazımdır. Belə qruplara anlayışlar deyilir. Bu sinif analiz problemləri oxşarlıq əlaqələrinin riyazi nəzəriyyəsinə əsaslanır.

Rəsmi konseptual analizin fərqli bir xüsusiyyəti ondan ibarətdir ki, onun köməyi ilə xüsusiyyətləri bənzər obyektlərin qruplarını avtomatik seçmək və aralarında əlaqələr qurmaq mümkündür. Konseptual oxşarlıq nəzəriyyəsi Ganter və Wille (1989) tərəfindən hazırlanmışdır. Konseptual bənzərlikdəki ümumi proses mövzu sahəsi haqqında biliklərin ixtiyari dəyərləri və bəlkə də itirilmiş dəyərləri olan məlumat cədvəlləri şəklində təqdim edilməsindən başlayır. Rəsmi konseptual təhlildə mahiyyəti MD($n \times d$) verilənlər bazası olan məlumat cədvəlləri xüsusi kontekst adlandırılır. [23]

Tərif 7.1. Rəsmi kontekstdə bu aşağıdakı üçlüdür:

$$C=(G, Y, I) \quad (2)$$

G -cisimlər toplusudur, atributlar toplusudur, $I \subseteq G \times Y$ isə G və Y arasındakı ikili əlaqədir. Üstəlik, $\langle g, y \rangle \in I$ ifadələrinin semantikasi belədir: "Cismin atributu var". (Gugisch, R. Many-valued Context Analysis using Descriptions)

Yəni, hər $y \in Y$ W üzərində bir xüsusiyyətdir (bəzi altçoxluqlarda $w \in W$). Bu o deməkdir ki, xüsusi kontekstdə $m \in M$ hər bir atributu üçün, $w \in W$ -nin mümkün qiymətləri çərçivəsində $w \in W$ -nin hər bir qiymətini $y \in Y$ -in bəzi xüsusiyyətlərinə çevirmək üçün interval, linqvistik termin və ya bir üzv funksiyası ilə qeyri-səlis aralıq vasitəsilə ifadə edilə bilən müəyyən ölçü qurmaq lazımdır.

Tərif 7.2. Konsepsiya (anlayış) bənzər xüsusiyyətləri ilə xarakterizə olunan $y \in Y$ obyektlərinin maksimum toplusudur (G obyektinin altçoxluğu), yəni oxşar xüsusiyyətlərə sahib olan obyektlərin qruplaşmasıdır.

Formal təhlilin vəzifəsi xüsusi kontekstin rəsmi bir kontekstə ilkin çevrilməsi əsasında anlayışları və münasibətləri müəyyənləşdirməkdir: $K \Rightarrow C$. Təhlil nəticəsində obyektləri ümumi xassələrə görə qruplaşdırmaq, ortaq xassələrə və qismən sifariş münasibətlərinə sahib olan obyektləri əks etdirən qrafik şəklində konseptual əlaqə yaratmaq mümkündür. Konseptual şəbəkəyə əsaslanaraq heterogen atributlara bənzər xüsusiyyətləri olan qruplar haqqında biliklər tapmaq, müxtəlif asılılıqlar tapmaq, iyerarxiya qurmaq, o cümlədən proqram mühəndisliyi sahəsində də biliklər əldə etmək mümkündür.

Rəsmi konseptual təhlil ədəbiyyatlarda geniş şəkildə təqdim olunmadığından, xüsusi və formal kontekstləri izah edən qısa bir nümunə vermək olar.

Xülasə və aqreqasiya vəzifələrinin qoyuluşu

Xülasə ayrı bir atribut və ya bu atributun xassəsi ilə əlaqəli xüsusi kontekst şəklində (məsələn, (2) şəklində) və rəsmi kontekst şəklində təmsil olunan bir sıra obyektlərin ümumi xarakteristikası olaraq təyin olunur. Xülasənin nəticəsi aqreqasiya əməliyyatının kəmiyyət xarakteristikası kimi ifadə edilə bilər. Məsələn, hər bir

klasterdəki obyektləri cəmləşdirmək, nümunənin orta, minimum, maksimum dəyərlərini və ya istədiyiniz çoxluqla əlaqəli məlumatların faizini hesablamaq.

Linqvistik xülasə kəmiyyət xüsusiyyətlərindən dil xüsusiyyətlərinə (dil qiymətləndirmələrinə), insan üçün başa düşülən və faydalı məlumatları ifadə etməyə imkan verir.

Tədqiqat obyektlərinin ayrı bir atribut (xassə) tərəfindən linqvistik ümumiləşdirilməsi üçün Yager tərəfindən təklif olunan əsas yanaşmadan istifadə etmək tövsiyə olunur. [29]

Əsas yanaşma çərçivəsində n obyektlər toplusu, y atributu ilə $D = \{y(g_1), \dots, y(g_n)\}$ verilənlər bazasında $G = \{g_1, \dots, g_n\}$ qeydləri şəklində, $y(g_i)$ isə işçi g_i üçün xüsusiyyətlər hesab olunur.

Xülasənin nəticəsi, həqiqət dərəcəsi olaraq modelləşdirilən bəzi qeyri-müəyyənlikləri nəzərə alaraq aşağıdakı komponentlərdən ibarətdir:

$g's$ - tədqiqat obyektinin (məsələn, tədqiqatçı) y atributu ilə əlaqəsinin təyin edilməsi, məsələn, $y = \text{"əmək haqqı"};$

P - xülasə nəticəsinin linqvistik mənası (məsələn, y atributu üçün "aşağı" = "əmək haqqı");

Q - tədqiq olunan obyektlərin P ilə əlaqəli kəmiyyət qiymətləndirməsinə əsaslanan linqvistik kəmiyyət (məsələn, "əksəriyyət", "bütövlükdə", "azlıq", "hamısı");

T - xülasə / ifadənin həqiqət dərəcəsidir.

Həqiqət dərəcəsi, tədqiqat obyektlərini atributları dəsti ilə avtomatik olaraq ümumiləşdirərkən ən əhəmiyyətli xarakteristikanı seçmək üçün lazımdır.

Beləliklə, linqvistik xülasənin əsasını mücərrəd formada yazıla bilən cümlələr məcmusu təşkil edir:

$Q\ g's\ P$ mənasına malikdir.

$Y = \text{"əmək haqqı"}$ atributu olan D verilənlər bazası üçün təklif nümunəsi aşağıdakı kimidir:

"İşçilərin ($g's$) az bir qismi (Q) yüksək (P) maaş alır (Y)".

Bu formatda linqvistik xülasə tətbiq edilərkən, Y atributu üçün üzv funksiyaları ilə qeyri-səlis linqvistik qiymətləndirmələr yaratmaq və bir çox obyektlərin bəzi kəmiyyət bölgüsü əsasında Q-nın keyfiyyətə qiymətləndirilməsi üçün dil ölçüsü hazırlamaq lazımdır.

Tərif 8.1. D verilənlər bazası üçün Y atributu və Q xarakteristikası ilə əlaqədar olaraq linqvistik xülasə olan P-nin nəticəsinin doğruluğu, Q obyektlər çoxluğunun Y atributunda P dəyərinə malik olduğu barədə qərarın düzgünlüyünü ifadə edir.

FƏSİL III. İNTELLEKTUAL TƏHLİL ÜSULLARININ TƏTBİQİNİN METODİKİ VƏ PRAKTİK ASPEKTLƏRİ

3.1. İqtisadi məlumatların intellektual təhlili üçün göstəricilər sisteminin qurulması

Dünyadakı ölkələrin dayanmadan inkişafı eyni zamanda ölkə iqtisadiyyatından dünya iqtisadiyyatına keçid tədqiqatçıları üçün yeni problemlər yaradır. Bir neçə problemdən biri inkişafın təhlükəsizliyi, dayanıqlılığı və ya davamlılığı problemidir. Bu məsələnin həlli üçün digər ölkələrin və təşkilatların statistikasının əsas konsepsiya və göstəricilərinin sistemləri dəqiqləşdirilməlidir.

Hal-hazırda tədqiqatçıların iqtisadi, demoqrafik, ekoloji təcrübələri, çətin və obyektiv həllərin tapılması üçün tematik həll daşıyıcısıdır, hansıki düzgün qərarlar üçün çox da əhəmiyyətli deyildir. Qeyd edə bilərik ki, istifadə etdiyimiz tədqiqat metodları dövrümüzün tələbatlarını qarşılamır. Elmin yeni tövhiyələrindən həmçinin informasiya texnologiyalarından yararlanmaqla inkişafın müxtəlif faktorlarının yəni, iqtisadi, demoqrafik, ekoloji, sosial dəyişikliyin dövlətin təhlükəsiz, stabil, kəsilməz inkişafına göstərdiyi təsirlərin öyrənilməsi öz əhəmiyyətini qoruyur.

BMT-nin 1992-ci ildə Rio-de-Janeyro da keçirdiyi beynəlxalq konfransda deyildiyi kimi ekoloji problemin həll metodlarını öyrənməklə dünyanın inkişaf artımını yüksəltmək olar. Keçirilən beynəlxalq konfransda ekoloji problemin həlli metodları məqsədli şəkildə müzakirə olunmuş və vacib qərarlar həmçinin bu məsələlər ilə məşğul olan sahələrin proqram sənədi olan – Dayanıqlı inkişaf ideyası qəbul olunmuşdur. Dayanıqlı inkişaf konsepsiyasında iqtisadi inkişafın ekoloji inkişafı ilə birlikdə yaratması, məqsədin ekoloji mühitin təhlükəsizliyinin təmin olunmasına yönləndirilməsi, ilk olaraq ekoloji məsələlərin həll olunması, növbəti nəsillərin imkanlarını azaltmadan, mövcud nəsillərin həyat şəraitinin inkişafı əsas olaraq qeyd edilmişdir. Bunu əldə etmək üçün ilk olaraq təbii ehtiyatlardan səmərəli

istifadəni yüksəltmək və ekoloji mühitin qorunmasını və bərpa olunmasını təmin etmək lazımdır.

Hal-hazırda ekoloji mühitlə əlaqədar olan məsələlərə ümumbəşəri kontekstdə yaxınlaşmaq lazımdır. Bununla əlaqədar olaraq ətraf mühitin təhlükəsizlik məsələləri əsas beynəlxalq qurumların proqram sənədlərində əks olunur.

Respublikamız son dövrlər sosial-iqtisadi inkişafına yönələn bir sıra nailiyyətlər qazanmışdır ki, bu da öz əksini beynəlxalq və milli sənədlərdə tapmışdır. Sosial-iqtisadi yöndə olan inkişafın davamlı olması ölkənin əsas məqsəd kimi qəbul olunmuşdur. Ölkədəki ekoloji siyasət strategiyası, ekoloji mühitin mühafizəsi yönündə fəaliyyətin koordinasiyasının artırılması yolu ilə beynəlxalq, regional, milli səviyyələr üzrə təbii ehtiyatların qorunması, elmə yönələn artım prinsipinin tətbiqi, indiki və sonrakı nəsillərin maraqlarını qarşılayan ölkənin iqtisadi davamlılığının təminatına yönəlmişdir. İqtisadi proseslərdə mühüm şərtlərdən biri əhalinin rifahı, ekologiyanın əhəmiyyəti və digər bu kimi faktorların bir-birilə qarşılıqlı əlaqəsi haqqında statistikaya tələbatdır. Bu statistikaya antiməhsul və məhsulun istehlakı və istehsalı statistikasına, təbii resursların statistikasına, tullantıların statistikasına, həyat keyfiyyəti və tullantıların udularaq zərərsizləşdirilməsi statistikasına və digər statistikalar daxildir.

Bilirik ki, xidmət və məhsulların istehsalı üçün təbii resurslar olmalıdır, lakin istehlak və istehsal təbii resursların azalması, tullantıların əmələ gəlməsi və davamlı şəkildə artmasına səbəb olmuşdur. Lakin bu zamanadək müxtəlif səbəblər nəticəsində təbii ehtiyatların azalması, tullantıların ətraf mühitə göstərdiyi mənfi təsirlər milli sərvətlərin azalmasına zəmin yaratmır çünki təbiətə təsir göstərən əsas faktor insan fəaliyyətidir və təbiət özünü əvvəlki vəziyyətinə qaytarmaq qabiliyyətinə malikdir. Əhalinin rifahına ətraf mühitin təsirinin olması böyük məlumat bazası və mürəkkəb metodoloji məsələlərin həllinin olmasını zəruri edir. Təbii resursların azalmasına, tükənməsinə səbəb insanların ətraf mühitə laqeyd yanaşması oldu ki, bunun da nəticəsində problem daha da global xarakter aldı.

İqtisadi, ekoloji və sosial dəyişikliklərin statistikasını yəni, ən yaxşı göstəricilər sistemi olmasa ətraf mühitə sosial və iqtisadi fəaliyyətin təsirinin öyrənilməsi və həmin təsirin aradan götürülməsi üsullarının tapılması indiki materialın hazırlanmasını vacib etmişdir.

Azərbaycan Respublikasının Prezidenti 2016-cı il 6 oktyabr tarixində “Azərbaycan Respublikası Dayanıqlı inkişaf üzrə Milli Əlaqələndirmə Şurası”nın yaradılması haqqında fərman imzalamışdır. Fərmanda qeyd olunmuşdur ki, Minilliyin Bəyannaməsinə qoşulan ölkəmiz yoxsulluqla mübarizə aparmaq üçün müxtəlif tədbirlər həyata keçirmək məqsədilə “2003-2005-ci illərdə Azərbaycan Respublikasında iqtisadi inkişaf və yoxsulluğun azaldılması istiqamətində Dövlət Proqramı” həmçinin “2008-2015-ci illər ərzində Azərbaycan Respublikasında davamlı inkişaf və yoxsulluğun azaldılması istiqamətində Dövlət Proqramı” işlənib hazırlanmış və icra olunmuşdur. Azərbaycanın son illər çərçivəsində Minilliyin inkişafı Məqsədi üzrə hədəfləri, xüsusilə də yoxsulluğun aradan götürülməsi, gender bərabərliyinin təmin olunması, əhəlinin təhsil sahəsində nəticələrinin yüksəldilməsi, insanların sağlamlığının və ətraf mühitin qorunması üzrə böyük nailiyyətlər qazanmışdır.

Respublikamız 2015-ci il 25-27 sentyabr tarixində BMT-yə üzv olan dövlətlərin keçirdiyi Dayanıqlı İnkişaf Sammitində təsdiqlənmiş 2016-2030-cu illər üçün nəzərdə tutulan Dayanıqlı İnkişaf Məqsədinə qoşulmuşdur. Qarşıya qoyulan hədəflər və məqsədlər dayanıqlı inkişafın sosial, ekoloji, iqtisadi aspektlərini əhatə etmək ilə yanaşı növbəti inkişaf istiqamətini müəyyən edir və Minilliyin İnkişaf Məqsədinə əsaslanır.

Dayanıqlı İnkişaf Məqsədinin icrası siyasi və institusional baxımdan kompleks yaxınlaşma tələb etdiyinə görə, Azərbaycanda daha şaxələndirilmiş, dayanıqlı və əhatəli iqtisadiyyatın, sosial sahələrin inkişaf etdirilməsi üçün xüsusi əhəmiyyət kəsb edən amil məqsədəuyğun milli icra mexanizminin formalaşmasıdır. Ona görə də “Dünyamızın dəyişməsi: 2030-cu ilə qədər davamlı inkişaf sahəsindəki Gündəlik”

dən yaranan öhdəliklərə görə Milli Əlaqələndirmə Şurası yaradılmış və bu Şuranın məqsədi dövlət orqanlarının vəzifələrinin icrasının əlaqələndirilməsidir.

Avropa İttifaqı Statistika Bürosu göstəriciləri əsasında Dayanıqlı inkişafın göstəriciləri hazırlanmışdır. Bu göstəricilər altbölmədən ibarət olub 10 mövzuya bölünür:

1. Dayanıqlı istehsal və istehlak;
2. Sosial-iqtisadi inkişaf;
3. Demografik dəyişikliklər;
4. Sosial tendensiyalar;
5. Enerji və iqlim dəyişiklikləri;
6. Səhiyyə;
7. Təbii resurslar;
8. Nəqliyyatın inkişafı;
9. Səmərəli idarəetmə;
10. Qlobal tərəfdaşlıq.

Sosial-iqtisadi göstəricilər: bir nəfər üçün nəzərdə tutulmuş regional idxal olunan məhsulun paylanması, əhali içərisindən bir nəfərin payına düşən ÜDM, ev təsərrüfatı yığımının inkişaf sürəti, birbaşa əldə olunan milli gəlir, real valyuta məzənnəsi, işləmələrə və tədqiqata sərf olunan xərclər, iqtisadiyyatın enerji həcmi, innovasiyaların dövrüyyəsi, işsizlik səviyyəsi, işgüzarlığın həcmi, tullantılar və onların zərərsiz hala salınması, materiala çəkilən xərclərin tərkibi, iqtisadi sahələr üzrə havay atılan kükürd oksidi-SO_x, sahələr üzrə havay atılan azot oksidi-NO_x, iqtisadi fəaliyyət tipləri üzrə təhlükəli olan lazımsız məhsulların əmələ gəlməsi, sahələr üzrə havaya atılan amonyak-NH₃, sahələr üzrə havaya atılan metal olmayan az çəkili üzvi birləşmələr, adambaşına düşən ərzaq istehlakı, enerjinin axırıncı istehlakı zamanı bərpa olan enerji miqdarı, aqroekoloji torpaqlar, avtomobil ilə təmin olunma səviyyəsi, kənd təsərrüfatı heyvanları üçün əldə olunan sıxlıq indeksi, orqanik əkinçiliyə görə istifadə edilən əkin sahəsi, məqsədlərə əsaslanan ev təsərrüfatının sonuncu istehlak xərcləri, ev təsərrüfatına daxil olan üzvlərin sayı.

Demoqrafik göstəricilər: ümumi nəsilartırma əmsalı, 65 yaşda ehtimal olunan uzunömürlülük, yaşlı təbəqənin məşğulluq səviyyəsi, ümumi yerdəyişmə əmsalı, əmək qabiliyyəti olan əhalinin uzunömürlülüüyü, gənc nəslin əvəz olunma səviyyəsi, yaşlı əhalinin asılılıq əmsalının proqnozlaşdırılması, yaşlı əhalinin asılılıq əmsalı, qocalara qulluq xərcləri, pensiya xərclərinin proqnozlaşdırılması.

Sosial göstəricilər: sosial yerdəyişmələr istisna olmaqla əhalinin cinslər üzrə yoxsulluğu, yoxsulluq riski içərisində olmanın daimiliyi, gəlirlərin bölüşdürülməsi, sosial yerdəyişmələr istisna olmaqla əhalinin yaş səviyyələri üzrə yoxsulluq, əhalinin cinslərinə görə uzunmüddət davam edən işsizlik, iqtisadi aktivliyi zəif olan ev təsərrüfatında yaşayan yoxsul təbəqə, ən yüksək təhsil pilləsinə görə yoxsulluq, əmək haqqı üzrə gender bərabərsizliyi, fasiləsiz təhsil, yaş qruplarına görə təhsili yüksək səviyyədə olmayan əhali, fiziki şəxslərin kompüterdən istifadə qabiliyyəti, oxumaq qabiliyyəti yüksək olmayan şagirdlər, dövlətin təhsilə sərf etdiyi xərclər, fiziki şəxsin internetdən istifadə qabiliyyəti.

Enerji göstəriciləri: bərpa oluna bilən mənbələr vasitəsilə elektrik enerjisinin yaranması, yanacaq növlərinə görə enerji istehlakı, elektrik və istilik enerjisinin birlikdə istehsalı, enerji varlığının dərəcəsi, nəqliyyatda istifadə olunan yanacağın tərkibindəki bərpa oluna bilən enerji növlərinin miqdarı.

İqlim dəyişikliyi göstəriciləri: zərərli qazların havaya atılmasının proqnozlaşdırılması, havanın normal temperaturu, enerji paylanması nəticəsində yaranan zərərli qazların havaya atılması intensivliyi.

Səhiyyə göstəriciləri: əhalinin cins üzrə xroniki xəstəlikdən ölüm halları, əhalinin cins üzrə uzunömürlülüüyü, əhalinin cinsinə və yaş qruplarına görə intihar hallarının sayı, əhalinin cins üzrə 65 və ondan yuxarı yaş həddi üçün uzunömürlülük, şəhər əhalisinə havanın bərk maddələrdən ibarət tullantılarla çirklənməsinin təsiri, gəlirlərin səviyyəsinə görə tibbi müayinə və müalicə tələbatının ödənilməsi, istehsalatda xoşagəlməz hadisələr, səs təsirindən narahat olan əhalinin miqdarı.

Təbii resurslar: quşların miqdarının dəyişmə indeksi, biomüxtəliflik, şəhərlər üzrə çirkli suların təmizlənməsi səviyyəsi, ümumi su ehtiyatları içərisində yeraltı və

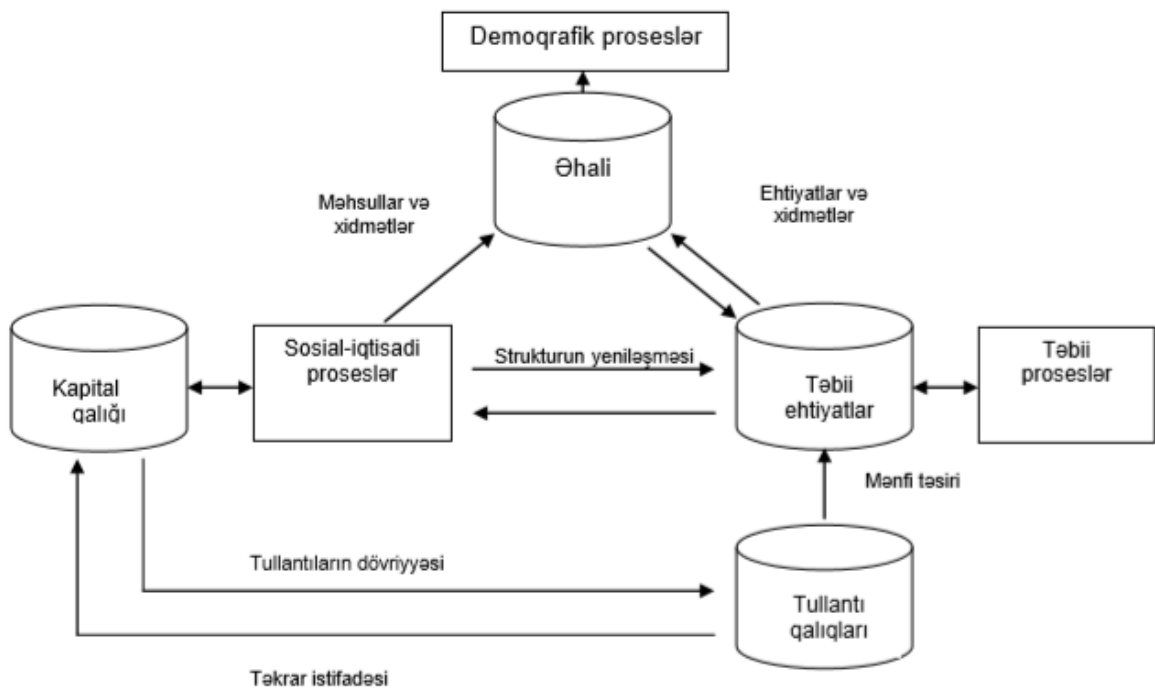
yerüstü suların payı, təhlükəli olmayan bioloji zonlardan uzaqda yerləşən ehtiyatdan balıq ovu, çayların oksigenə duyduğu biokimyəvi tələbat, meşələrin məhv edilməsi və bərpa, balıqçılıq donanmasının tutumu.

Nəqliyyat göstəriciləri: sənişin daşınmasının miqdarının ÜDM-ə nisbəti, yükdaşımaların miqdarının ÜDM-ə nisbəti, nəqliyyat növlərinə əsaslanan nəqliyyat infrastrukturuna olan investisiyalar, nəqliyyat növlərinə əsaslanan enerji istehlakı, nəqliyyat vasitələrindən ətraf mühitə buraxılan bərk hissəciklərin həcmi, nəqliyyat vasitələrindən ətraf mühitə buraxılan azot oksidin həcmi, nəqliyyat xidmətinə əsasən tariflərin qiymət əmsalı, yeni avtomobillər vasitəsilə havaya buraxılan karbon qazının miqdarı.

Səmərəli idarəetmə: vətəndaşların idarəetmə institutlarına inam səviyyəsi, ekoloji, sosial və əmək ödəmələrinin ümumi ödəmələrdəki payı, elektron hökumətdən fiziki şəxslərin istifadəsi, büdcə kəsiri.

Qlobal tərəfdaşlıq: əhalinin bir nəfəri üçün nəzərdə tutulmuş havaya buraxılmış karbon qazının həcmi, əhalinin bir nəfəri üçün nəzərdə tutulan rəsmi qəbul edilmiş yardımların məbləği, inkişaf məqsədlərinə görə ikitərəfli yardımlar, ölkələrin gəlir qruplarına əsasən yardımları, birbaşa investisiyalar, inkişaf etməmiş ölkələrə yardımların həcmi, kənd təsərrüfatı üçün nəzərdə tutulan subsidiyaların həcmi, inkişaf etmiş ölkələr üçün mal qruplarına görə ixrac, inkişaf etmiş ölkələrdən mal qruplarına görə idxal, inkişaf üçün nəzərdə tutulan rəsmi yardımların ÜMG-də payı.

Sosial-iqtisadi, ekoloji, demoqrafik ünsürlərin qarşılıqlı təsiri sxemdə aşağıdakı şəkildə göstərilmişdir:



Ölkəmizin daimi inkişafına əsaslanan aşağıdakı göstəricilərin tədqiqatını statistik metodlar vasitəsilə aparaq. Aşağıdakı şərti işarələr vasitəsilə tədqiqatı apara bilərik:

Y_1 - ÜDM (milyon AZN);

Y_2 - Adambaşına düşən ÜDM (AZN);

Y_3 - Adambaşına düşən enerji istehlakı;

Y_4 - ÜDM-də investisiyanın miqdarı (%);

Y_5 - Əhalinin təbii şəkildə artımı (hər 1000 nəfər üçün);

Y_6 - Ətraf mühitin qorunmasına sərf olunmuş xərclərin ÜDM-dəki payı;

Y_7 - Doğum səviyyəsi;

Y_8 - Əhalinin ömür uzunluğu;

Y_9 - Uşaqların ölüm səviyyəsi (1-4 yaş aralığında);

Y_{10} - Körpələrin ölüm səviyyəsi (1 yaşadək);

Y_{11} - İlin sonunda şəhər əhalisi (%);

Y_{12} - İşsizlik səviyyəsi;

Y_{13} - Əhalinin sıxlığı;

Y_{14} - Əhalinin hər bir nəfərinə düşən mənzilin ümumi sahəsi;

Y_{15} - Adambaşına düşən əkin sahəsi;

Y_{16} - Adambaşına düşən su istehlakı;

Y_{17} - Adambaşına düşən bərk məişət tullantısı;

Y_{18} - Kənd təsərrüfatı sahəsində enerji istehlakı;

Y_{19} - Stasionar mənbələrdən havaya atılan adambaşına çirkləndirici maddələr;

Y_{20} - Stasionar mənbələrdən havaya atılan adambaşına qazaoxşar və maye maddələr.

Dayanıqlı inkişafı xarakterizə edən göstəricilərin dinamikası

	2007	2008	2009	2010	2011	2012	2013	2014	2015
Y_1	28360,5	40137,2	35601,5	42465	52082	54743,7	58182	59014,1	54352,1
Y_2	3296,6	4603,7	4033,2	4753	5752,9	5966,1	6258,3	6268	5703,7
Y_3	1,5	1,7	1,5	1,4	1,5	1,6	1,6	1,6	1,6
Y_4	26,3	24,8	21,7	23,3	24,6	28,1	30,7	29,9	29,4
Y_5	11,5	11,4	11,3	12,5	13,5	13	12,8	12,2	11,7
Y_6	0,4	0,4	0,5	0,6	0,6	0,8	0,7	0,5	0,3
Y_7	17,7	17,4	17,2	18,5	19,4	19	18,6	18,1	17,4
Y_8	73	73,4	73,5	73,6	73,8	73,9	74,2	74,2	75,2
Y_9	16,1	14,9	14,4	14,2	13,5	13	12,9	12,6	13,3
Y_{10}	12,1	11,4	11,3	11,2	11	10,8	10,8	10,2	11
Y_{11}	53	53,1	53,1	53	52,9	53,1	53,2	53,1	53,1
Y_{12}	6,3	5,9	5,7	5,6	5,4	5,2	5	4,9	5
Y_{13}	101	103	104	105	107	108	109	111	112
Y_{14}	17,2	17,6	17,7	17,7	17,8	17,8	17,8	17,8	17,9
Y_{15}	0,154	0,172	0,193	0,177	0,178	0,18	0,181	0,171	0,166
Y_{16}	973	904,5	865,4	863,5	885	899	885,1	861,9	899

Y ₁₇	187	167,3	179,1	177,7	194,1	177,2	177,4	142,1	159
Y ₁₈	260,5	312,5	391	406,6	428	452,8	469,4	461,8	449,2
Y ₁₉	44	33	34	24	24	24	21	20	19
Y ₂₀	41	30	32	22	23	23	20	19	18

Dayanıqlı inkişafın əsas göstəricilərinin təsviri statistikası

	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅	Y ₆	Y ₇	Y ₈	Y ₉	Y ₁₀
Orta qiymət	47215,3	5181,7	1,6	26,5	12,2	0,5	18,1	73,9	13,9	11,1
Orta qiymətin standart xətası	3639,7	352,4	0,0	1,1	0,3	0,1	0,3	0,2	0,4	0,2
Moda			1,6			0,4	17,4	74,2		11,0
Median	52082,0	5703,7	1,6	26,3	12,2	0,5	18,1	73,8	13,5	11,0
Dispersiya	119228045,4	1117959,2	0,0	10,0	0,6	0,0	0,6	0,4	1,3	0,3
Standart kənarlaşdırma	10919,2	1057,3	0,1	3,2	0,8	0,2	0,8	0,6	1,1	0,5
Eksess	-0,6	-0,7	-0,2	-0,1	0,4	0,3	0,3	1,0	0,9	0,4
Asimmetriya	-1,0	-0,8	0,1	-1,4	-1,3	-0,5	-1,3	1,9	0,4	1,7
Ranq	30653,6	2971,4	0,3	9,0	2,2	0,5	2,2	2,2	3,5	1,9
Maximum	59014,1	6268,0	1,7	30,7	13,5	0,8	19,4	75,2	16,1	12,1
Minimum	28360,5	3296,6	1,4	21,7	11,3	0,3	17,2	73,0	12,6	10,2
Sayı	9,0	9,0	9,0	9,0	9,0	9,0	9,0	9,0	9,0	9,0
Cəm	424938,1	46635,5	14,0	238,8	109,9	4,8	163,3	664,8	124,9	99,8
Etibarlılıq səviyyəsi-95,0%	8393,2	812,7	0,1	2,4	0,6	0,1	0,6	0,5	0,9	0,4

	Y ₁₁	Y ₁₂	Y ₁₃	Y ₁₄	Y ₁₅	Y ₁₆	Y ₁₇	Y ₁₈	Y ₁₉	Y ₂₀
Orta qiymət	53,1	5,4	106,7	17,7	0,2	892,9	173,4	403,5	27,0	25,3
Orta qiymətin standart xətası	0,0	0,2	1,2	0,1	0,0	11,4	5,2	24,1	2,8	2,5
Moda	53,1	5,0		17,8		899,0			24,0	23,0
Median	53,1	5,4	107,0	17,8	0,2	885,1	177,4	428,0	24,0	23,0
Dispersiya	0,0	0,2	13,8	0,0	0,0	1166,2	239,9	5219,6	68,8	57,0
Standart kənarlaşdırma	0,1	0,5	3,7	0,2	0,0	34,2	15,5	72,2	8,3	7,5
Eksess	-0,7	0,6	0,0	-2,1	-0,4	1,8	-1,0	-1,3	1,2	1,3
Asimmetriya	0,8	-0,5	-1,1	5,0	1,3	4,0	1,1	0,6	0,8	1,0
Ranq	0,3	1,4	11,0	0,7	0,0	111,1	52,0	208,9	25,0	23,0
Maximum	53,2	6,3	112,0	17,9	0,2	973,0	194,1	469,4	44,0	41,0
Minimum	52,9	4,9	101,0	17,2	0,2	861,9	142,1	260,5	19,0	18,0
Sayı	9,0	9,0	9,0	9,0	9,0	9,0	9,0	9,0	9,0	9,0
Cəm	477,6	49,0	960,0	159,3	1,6	8036,4	1560,9	3631,8	243,0	228,0
Etibarlılıq səviyyəsi-95,0%	0,1	0,4	2,9	0,2	0,0	26,3	11,9	55,5	6,4	5,8

Təsviri statistikadan istifadə edərək əsas göstəricilər üçün dayanıqlılıq əmsalını, orta kəmiyyət xətasını və variasiyanı tapmaq olar.

Göstəricinin adı	Orta kəmiyyətin standart xətası, %-lə $\sigma / \sqrt{n}$	Variasiya əmsalı, faizlə $\sigma / \bar{x}$	Dayanıqlılıq əmsalı, faizlə $1 - (\sigma / \bar{x})$
Adambaşına ÜDM (AZN)	6,80	20,41	79,59
ÜDM (milyon AZN)	7,71	23,13	76,87
ÜDM-də investisiyanın payı (%)	3,97	11,92	88,08
Adambaşına enerji istehlakı, NET/nəfər	1,89	5,67	94,33
Ətraf mühitin mühafizəsinə çəkilən xərclərin ÜDM-də payı (%)	9,88	29,65	70,35
Əhalinin təbii artımı (hər 1000 nəfərə)	2,15	6,45	93,55
Əhalinin ömür uzunluğu	0,28	0,85	99,15
Doğum səviyyəsi	1,43	4,29	95,71
Körpə ölüm səviyyəsi (1 yaşadək hər 1000 nəfər diri doğulana)	1,56	4,67	95,33
Uşaq ölüm səviyyəsi (1-4 yaşda hər 1000 nəfər diri doğulana)	2,71	8,13	91,87
İşsizlik səviyyəsi (%)	2,89	8,67	91,33
Şəhər əhalisi ilin sonuna (%)	0,05	0,16	99,84
Əhalinin hər nəfərinə düşən mənzil sahəsi (m ² ümumi sahə)	0,39	1,16	98,84
Əhalinin sıxlığı	1,16	3,48	96,52
Adambaşına su istehlakı	1,27	3,82	96,18
Adambaşına əkin sahəsi (ha)	2,07	6,22	93,78
Kənd təs.-da enerji istehlakı, min NET	5,97	17,90	82,10
Adambaşına bərk məişət tullantıları (kg)	2,98	8,93	91,07
Stasionar mənbələrdən atmosfer havasına atılmış adambaşına düşən maye və qazaoxşar maddələr (kg)	9,93	29,80	70,20
Stasionar mənbələrdən atmosfer havasına atılmış adambaşına düşən çirkləndirici maddələr (kg)	10,24	30,71	69,29

Cədvəldən göründüyü qədər adambaşına düşən ÜDM (Y_2), kənd təsərrüfatı sahəsində enerji istehlakı (Y_{18}), stasionar mənbələrdən havaya atılan adambaşına çirkləndirici maddələr (Y_{19}), stasionar mənbələrdən havaya atılan adambaşına qazaoxşar və maye maddələr (Y_{20}) üzrə orta qiymətinin xətası 5 faizdən çox, həmin göstəricilərin variasiya əmsalı 10 faizdən çox, dayanıqlılıq əmsalı isə 80 faizdən çox olmuşdur.

Göstəricilərin normallaşdırılması üçün aşağıdakı düsturdan istifadə olunur:

$$Z = (x_i - \bar{x}) / \sigma$$

Dayanıqlı inkişafın göstəricilərinin normallaşdırılmış qiyməti

	2007	2008	2009	2010	2011	2012	2013	2014	2015	$\bar{x}$	σ
Y ₁	-1,73	-0,65	-1,06	-0,44	0,45	0,69	1	1,08	0,65	0	1
Y ₂	-1,78	-0,55	-1,09	-0,41	0,54	0,74	1,02	1,03	0,49	0	1
Y ₃	-0,63	1,64	-0,63	-1,76	-0,63	0,5	0,5	0,5	0,5	0	1
Y ₄	-0,07	-0,55	-1,53	-1,02	-0,61	0,5	1,32	1,06	0,91	0	1
Y ₅	-0,9	-1,03	-1,16	0,37	1,64	1	0,75	-0,01	-0,65	0	1
Y ₆	-0,84	-0,84	-0,21	0,42	0,42	1,69	1,05	-0,21	-1,48	0	1
Y ₇	-0,57	-0,96	-1,21	0,46	1,61	1,1	0,59	-0,06	-0,96	0	1
Y ₈	-1,37	-0,74	-0,58	-0,42	-0,11	0,05	0,53	0,53	2,11	0	1
Y ₉	1,97	0,91	0,46	0,29	-0,33	-0,78	-0,87	-1,13	-0,51	0	1
Y ₁₀	1,95	0,6	0,41	0,21	-0,17	-0,56	-0,56	-1,72	-0,17	0	1
Y ₁₁	-0,77	0,38	0,38	-0,77	-1,92	0,38	1,54	0,38	0,38	0	1
Y ₁₂	1,81	0,97	0,54	0,33	-0,09	-0,52	-0,94	-1,15	-0,94	0	1
Y ₁₃	-1,53	-0,99	-0,72	-0,45	0,09	0,36	0,63	1,17	1,44	0	1
Y ₁₄	-2,43	-0,49	0	0	0,49	0,49	0,49	0,49	0,97	0	1
Y ₁₅	-1,9	-0,25	1,69	0,21	0,31	0,49	0,58	-0,34	-0,8	0	1
Y ₁₆	2,34	0,34	-0,81	-0,86	-0,23	0,18	-0,23	-0,91	0,18	0	1
Y ₁₇	0,88	-0,4	0,37	0,28	1,33	0,24	0,26	-2,02	-0,93	0	1
Y ₁₈	-1,98	-1,26	-0,17	0,04	0,34	0,68	0,91	0,81	0,63	0	1
Y ₁₉	2,05	0,72	0,84	-0,36	-0,36	-0,36	-0,72	-0,84	-0,96	0	1
Y ₂₀	2,08	0,62	0,88	-0,44	-0,31	-0,31	-0,71	-0,84	-0,97	0	1

Cədvəldən də gördüyümüz qədər normallaşdırılmış göstəricinin standart kənarlaşması vahidə, orta qiyməti isə sıfıra bərabərdir.

Göstəricilərin normallaşdırılmış qiymətlərinin korrelyasiya əmsalı

	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅	Y ₆	Y ₇	Y ₈	Y ₉	Y ₁₀
Y ₁	1.0	1.0								
Y ₂		1.0								
Y ₃	0.4	0.4	1.0	0.5						
Y ₄	0.7	0.7		1.0						
Y ₅	0.6	0.7	-0.2	0.2	1.0	0.8				
Y ₆	0.4	0.5	-0.1	0.1		1.0				

Y ₇	0.5	0.6	-0.2	0.2	1.0	0.8	1.0	0.0		
Y ₈	0.8	0.7	0.3	0.6	0.2	-0.1		1.0		
Y ₉	-1.0	-1.0	-0.2	-0.6	-0.6	-0.5	-0.5	-0.7	1.0	1.0
Y ₁₀	-0.9	-0.9	-0.3	-0.5	-0.5	-0.4	-0.4	-0.6		1.0

	Y ₁₁	Y ₁₂	Y ₁₃	Y ₁₄	Y ₁₅	Y ₁₆	Y ₁₇	Y ₁₈	Y ₁₉	Y ₂₀
Y ₁₁		-0.4								
Y ₁₂	1.0	1.0								
Y ₁₃	0.4	-1.0	1.0	0.8						
Y ₁₄	0.3	-0.9		1.0						
Y ₁₅	0.2	-0.3	0.1	0.6	1.0	-0.8				
Y ₁₆	-0.2	0.6	-0.5	-0.8		1.0				
Y ₁₇	-0.5	0.5	-0.6	-0.4	0.2	0.3	1.0	-0.3		
Y ₁₈	0.3	-1.0	0.9	0.9	0.5	-0.7		1.0		
Y ₁₉	-0.2	0.9	-0.9	-0.9	-0.3	0.7	0.4	-0.9	1.0	1.0
Y ₂₀	-0.2	0.9	-0.9	-0.9	-0.3	0.7	0.5	-0.9		1.0

Əhalinin təbii artımı (0.7), ÜDM-də investisiyanın miqdarı (0.7), doğum səviyyələri (0.6), əhalinin ömür uzunluğu (0.7), uşaqların ölüm səviyyəsi (-1), körpələrin ölüm səviyyəsi (-0.6), əhalinin hər bir nəfərinə düşən mənzilin ümumi sahəsi (0.8), işsizlik səviyyəsi (-0.9), kənd təsərrüfatı sahəsində enerji istehlakı (-0.9), əhalinin sıxlığı (0.9), stasionar mənbələrdən havaya atılan adambaşına çirkləndirici maddələr (-0.9), stasionar mənbələrdən havaya atılan adambaşına qazaoxşar və maye maddələr (-0.9) göstəricilərinin adambaşına düşən ÜDM göstəricisi ilə korrelyasiya əlaqəsinin mütləq qiyməti 0.5-dən çox olmuşdur.

Əhalinin ömür uzunluğu (0.8), əhalinin təbii artımı (0.6), uşaqların ölüm səviyyəsi (-1), körpələrin ölüm səviyyəsi (-0.9), əhalinin hər bir nəfərinə düşən mənzilin ümumi sahəsi (0.8), işsizlik səviyyəsi (-1), kənd təsərrüfatı sahəsində enerji istehlakı (0.9), əhalinin sıxlığı (0.9), stasionar mənbələrdən havaya atılan adambaşına çirkləndirici maddələr (-0.9), stasionar mənbələrdən havaya atılan adambaşına qazaoxşar və maye maddələr (-0.9) göstəricilərilə ÜDM-də investisiyanın miqdarı göstəricisinin korrelyasiyasının mütləq qiyməti 0.5-dən yüksək olmuşdur.

Adambaşına düşən enerji istehlakı göstəricisi ilə ilin sonunda şəhər əhalisi göstəricisinin korrelyasiya əmsalı 0.6 olmuşdur.

Ətraf mühitin qorunmasına sərf olunmuş xərclərin ÜDM-dəki payı (0.6), doğum səviyyəsi (1) göstəricilərinin əhalinin artımı göstəricisi ilə korrelyasiya əmsalının mütləq qiyməti 0.5-dən yüksək olmuşdur.

Uşaqların ölüm səviyyəsi (-0.7), körpələrin ölüm səviyyəsi (-0.6), əhalinin hər bir nəfərinə düşən mənzilin ümumi sahəsi (0.8), işsizlik səviyyəsi (-0.8), kənd təsərrüfatı sahəsində enerji istehlakı (0.7), əhalinin sıxlığı (0.9), stasionar mənbələrdən havaya atılan adambaşına çirkləndirici maddələr (-0.8), stasionar mənbələrdən havaya atılan adambaşına qazaoxşar və maye maddələr (-0.8) göstəricilərinin əhalinin ömür uzunluğu göstəricisi ilə korrelyasiya əmsalının mütləq qiyməti 0.5-dən yüksək olmuşdur.

Əhalinin hər bir nəfərinə düşən mənzilin ümumi sahəsi (-0.8), işsizlik səviyyəsi (0.9), adambaşına düşən su istehlakı (0.7), əhalinin sıxlığı (-0.8), adambaşına düşən bərk məişət tullantısı (0.6), kənd təsərrüfatı sahəsində enerji istehlakı (-0.9), stasionar mənbələrdən havaya atılan adambaşına çirkləndirici maddələr (0.9), stasionar mənbələrdən havaya atılan adambaşına qazaoxşar və maye maddələr (0.9) göstəricilərinin körpələrin ölüm səviyyəsi ilə korrelyasiya əmsalının mütləq qiyməti 0.5-dən yüksək olmuşdur.

Adambaşına düşən su istehlakı (-0.9), əhalinin hər bir nəfərinə düşən mənzilin ümumi sahəsi (1), kənd təsərrüfatı sahəsində enerji istehlakı (-1), ilin sonunda şəhər əhalisi (-0.9), stasionar mənbələrdən havaya atılan adambaşına çirkləndirici maddələr (0.9), stasionar mənbələrdən havaya atılan adambaşına qazaoxşar və maye maddələr (0.9) göstəricilərinin uşaqların ölüm səviyyəsi göstəricisi ilə korrelyasiya əmsalının mütləq qiyməti 0.5-dən yüksək olmuşdur.

Kənd təsərrüfatı sahəsində enerji istehlakı (0.9), adambaşına düşən bərk məişət tullantısı (-0.8), stasionar mənbələrdən havaya atılan adambaşına çirkləndirici maddələr (0.9), stasionar mənbələrdən havaya atılan adambaşına qazaoxşar və maye

maddələr (0.9) göstəricilərinin əhalinin sıxlığı göstəricisi ilə korrelyasiya əmsalının mütləq qiyməti 0.5-dən yüksək olmuşdur.

Adambaşına düşən əkin sahəsi (0.8), kənd təsərrüfatı sahəsində enerji istehlakı (-0.7), stasionar mənbələrdən havaya atılan adambaşına çirkləndirici maddələr (0.7), stasionar mənbələrdən havaya atılan adambaşına qazaoxşar və maye maddələr (0.7) göstəricilərinin adambaşına düşən su istehlakı göstəricisi ilə korrelyasiya əmsalının mütləq qiyməti 0.5-dən yüksək olmuşdur.

Stasionar mənbələrdən havaya atılan adambaşına çirkləndirici maddələr (-0.9), stasionar mənbələrdən havaya atılan adambaşına qazaoxşar və maye maddələr (-0.9) göstəricilərinin kənd təsərrüfatı sahəsində enerji istehlakı göstəricisi ilə korrelyasiya əmsalının mütləq qiyməti 0.5-dən yüksək olmuşdur.

Faktor təhlili vasitəsilə yuxarıda verilən göstəricilərin dispersiya yükünü hesablamaq mümkündür. Y_1 -in dispersiya yükü 16.97%, Y_2 -nin dispersiya yükü 57.33%, Y_4 -ün dispersiya yükü 12.79%, 3 komponentin birlikdə dispersiya yükü 87.09% təşkil edir. Digər 13 göstəricinin dispersiya yükü 13%-dir.

Rotasiya vasitəsilə göstəriciləri 3 faktora bölüb, bu faktorların verilən göstəricilər ilə korrelyasiya indeksini aşağıdakı şəkildə vermək olar: [33]

	1	2	3
Y_1	0.983		
Y_2	0.967		
Y_3			
Y_4	0.747		0.579
Y_5		0.864	
Y_6		0.719	
Y_7		0.922	
Y_8	0.846		
Y_9	-0.956		
Y_{10}	-0.906		
Y_{11}	-0.982	-0.566	
Y_{12}			

Y ₁₃	0.965		
Y ₁₄	0.846		
Y ₁₅			-0.898
Y ₁₆	-0.515		0.817
Y ₁₇	-0.561	0.685	
Y ₁₈	0.901		
Y ₁₉	-0.936		
Y ₂₀	-0.927		

Birinci faktor sosial-iqtisadi faktordur və onun ÜDM (0.983), adambaşına düşən ÜDM (0.967), əhalinin ömür uzunluğu (0.846), ÜDM-də investisiyanın miqdarı (0.747), uşaqların ölüm səviyyəsi (-0.956), körpələrin ölüm səviyyəsi (-0.906), əhalinin hər bir nəfərinə düşən mənzilin ümumi sahəsi (0.846), əhalinin sıxlığı (0.965), işsizlik səviyyəsi (-0.982), adambaşına bərk məişət tullantısı (-0.561), kənd təsərrüfatı sahəsində enerji istehlakı (0.901), stasionar mənbələrdən havaya atılan adambaşına çirkləndirici maddələr (-0.936), stasionar mənbələrdən havaya atılan adambaşına qazaoxşar və maye maddələr (-0.926) göstəricilərlə korrelyasiyası yüksək olmuşdur. İkinci faktor demoqrafik-ekoloji faktordur və onun əhalinin təbii artımı (0.864), ətraf mühitin qorunmasına çəkilən xərclər (-0.568), kənd təsərrüfatı sahəsində enerji istehlakı (-0.568), adambaşına düşən bərk məişət tullantısı (0.685), doğum səviyyəsi (0.926) göstəricilərlə korrelyasiya yüksək olmuşdur. Üçüncü faktor isə maliyyə faktorudur və onun ÜDM-də investisiya miqdarı (0.579) göstəricisi ilə korrelyasiya yüksək olmuşdur. Korrelyasiya əmsalının qiyməti müsbət olduqda əlaqə düz olur yəni arqumentin artması faktorun artmasına səbəb olur. Korrelyasiya əmsalının qiyməti mənfi olduqda əlaqə tərs olur yəni arqumentin artması faktorun azalmasına səbəb olur.

Dayanıqlı inkişafı müəyyən edən göstəricilərin təhlilindən məlum olur ki, ÜDM istehsal göstəricisi olduğuna görə mövcud olan sosial-ekoloji və sosial-iqtisadi münasibətləri hərtərəfli xarakterizə edə bilmir. Ona görə də əksər tədqiqatçılar bu göstəricini Xalis Milli Gəlir, bəziləri isə Xalis Daxili Məhsul göstəricisi ilə əvəz olunmasını tələb edir.

Postsovet dövlətləri müstəqillik qazandıqdan sonra insan inkişafı və iqtisadi indeks hesablarında statistiklərin qarşısındakı vacib tədqiqat məsələlərindən biri ekoloji hesabatların təşkili, ekoloji inkişaf indeksinin hesablanması və bu hesabların iqtisadi hesablar ilə inteqrasiyası olmuşdur.

Müəyyən edilmişdir ki, yaranan ekoloji problemlər istehsalçılar və istehlakçılar yəni iqtisadçılar üçün yeni məsələlər yaradır. Bu məsələlərdən biri isə ÜDM-də ekoloji məhsulun minimumlaşdırılması və ekoloji məhsuldan səmərəli istifadədir. İqtisadiyyatın davamlı şəkildə inkişafı üçün təbii ehtiyatlara qənaətlə yanaşılmalı, yerüstü və yeraltı sərvətlər mühafizə olunmalıdır. Əgər bu məsələlərə əməl olunmazsa iqtisadiyyatın davamlı inkişafı mümkün olmaz. Hətta ekoloji məhsul yəni enerji və ərzaq daşıyıcıları da tükənər. Bazar iqtisadiyyatı mühitində düzgün siyasətin qurulması üçün əsas amillərdən biri ekoloji məhsulların tələbi və təklifi (son istehlakı, aralıq istehlakı, ixracı, idxalı, istehsalı) göstəricilərin keyfiyyətli statistikasıdır.

Dövlət tərəfindən əhali artdıqca sosial-iqtisadi vəziyyətin sabilliyini qorumaq üçün enerji və ərzaq daşıyıcılarına subsidiyalar qoyulur. Lakin görülməli tədbirlər enerji və ərzaq daşıyıcılarının tükənməsinə və israfçılığına səbəb olur. Bazar iqtisadiyyatı mühitində insan və ekoloji kapitala nisbətən iqtisadi kapitala üstünlük verilir. Ona görə də bazar iqtisadiyyatında mənfi təzadları görmək lazımdır. Əgər bazar iqtisadiyyatı ekoloji həqiqəti deyə bilməzsə onda ətraf mühiti və özünü məhv edəcəkdir.

Davamlı inkişaf aşağıdakı model şəklində xarakterizə olunur:

$$F_t(E, K, L, I) \leq F_{t+1}(E, K, L, I)$$

E – təbii (ekoloji) kapital;

I – institusional faktor;

L – insan kapitalı;

K – iqtisadi kapitaldır.

Modeldən də göründüyü kimi insan kapitalı və təbii kapital daimi inkişafın bazis faktorudur yəni məhsuldar qüvvədir.

Ona görə də məhsuldar qüvvələrin qorunub saxlanması, bərpası və onlardan səmərəli istifadə strateji məqsəd kimi qalmaqdadır. Bu strategiyayı reallaşdırmaq üçün aşağıdakı şərtlər ödənilməlidir:

- a) istehsal ekologiyalaşdırılmalı;
- b) istehlak məqsədəuyğun olmalı;
- c) ölkədəki son istehlak təbii məhsuldarlığa uyğunlaşmalı;
- d) ölkənin ekoloji təhlükəsizliyi üçün siyasət hazırlanmalı;
- e) əhalinin sayı optimallaşdırılmalıdır.

3.2. İqtisadi göstəricilərin intellektual təhlili (DM) üsulları ilə

proqnozlaşdırılması

Ölkənin uğurlu iqtisadi inkişafı üçün iqtisadiyyatın real sektorunun, ilk növbədə sənayenin inkişaf etdirilməsinə yönəlmiş dövlət siyasəti lazımdır.

"Yeni İqtisadi Sözlük" də sənayeyə xalq təsərrüfatının sektoru deyilir, hansiki sənaye mallarının istehsalı və istehsal xidmətlərinin göstərilməsi ilə əlaqələndirilir. Xalq təsərrüfatının sənaye sektoru dağ-mədən, yüngül, kiçik, istehsal, ağır və aksiz sənayəsindən ibarətdir. “ Böyük İqtisadi Sözlük “ də deyilirki sənaye - sənaye mallarının istehsalı və istehsal xidmətlərinin göstərilməsi ilə əlaqəli iqtisadiyyatdır.

Bu gün ölkədə mövcud olan bütün sənaye ehtiyatlarından effektiv və səmərəli istifadə məsələsi kimi xammal hasilatından tutmuş materialların, enerjinin istehsalına

və emalına, ardınca həm daxili, həm də beynəlxalq bazarlarda tələb olunan maşın, texnikası, avadanlıq və istehlak mallarının istehsalına baxılır.

Problemin ifadəsi iqtisadi inkişafın təhlili və dövlətin əsas sənaye göstəricilərinin proqnozlaşdırılmasıdır. Maraqlı nəticələr mövqeyə görə iqtisadi inkişafın nisbi xüsusiyyətlərinin birincisi ümumi məhsul vahidinə görə cari xərclərin təhlilindən (vahid xərclərinin xüsusi göstəricisi), ikincisi cari xərclərin ümumi əlavə dəyərə əvəz edilməsindən yaranır (xərc səmərəliliyinin müəyyən bir göstəricisi).

Maraqlı nəticələr, iqtisadi inkişafın nisbi xüsusiyyətlərinin mövqeyə, ilk növbədə ümumi məhsul vahidinə görə cari xərclərin təhlilindən irəli gəlir.

$$W = 3626 + 5952t - 1459,4t^2 + 108,7t^3, \quad R^2 = 0,800, \quad F = 6,7$$

Hesablanmış model reallığı yaxşı təsvir edir (korrelyasiya əmsalı 0.894, müəyyənləşdirmə əmsalı isə 0.800), bu da təhlil olunan göstəricinin zaman amili ilə çox yaxın əlaqəsini göstərir. [32]

Sənayə inkişafını proqnozlaşdırmaq üçün müəyyən hesablamalara əsasən arzu olunan proqnoz göstəricilərinin (ümumi dəyər) onun amillərindən asılılığı aşkar edilməli və amil analizinə əsaslanan metodoloji bir yanaşmadan istifadə olunmalıdır. Beləliklə, aşağıdakı ardıcılıqdan istifadə edilməlidir:

- aralıq istehlak həcmələri üçün vaxt amili;
- ümumi məhsul üçün aralıq istehlak;
- ümumi əlavə dəyər üçün ümumi məhsul.

Tədqiqatın məqsədi dövlətin sosial-iqtisadi inkişafı göstəricilərinin proqnozunu əldə etməkdir. Daxil olan məlumatların və proqnozların təhlili prosesi nəzərdən keçirilir, prosesin mərhələləri (tədqiqatın məqsədinin modeldə iştirak edən izahlı dəyişənlərin məcmusundan formalaşdırılması; zəruri statistik məlumatların toplanması; seçilmiş proqnozlaşdırma texnikasına əsasən proqnoz göstəricilərinin hesablanması; analiz və proqnoz hesabatlarının generasiyası və vizuallaşdırılması), ilkin məlumatlara olan tələblər və sosial-iqtisadi proqnozların hazırlanmasında ən çox istifadə olunan metodlar təsvir edilir. Qeyd olunur ki, göstəricilərin ziddiyyətsiz və ardıcıl proqnozlarını əldə etmək üçün proqnoz göstəriciləri arasındakı qarşılıqlı

əlaqəni nəzərə almaq lazımdır. Deductor analitik platforması əsasında analitik tərəfindən təyin edilmiş skriptin icrası və hesabatların hazırlanması üçün informasiya-analitik sistemi hazırlanmışdır. Sistem məlumat toplusuna birləşdirilmiş bütün proqnozları hesablayır və minimum səhv ilə model haqqında məlumatlara əsaslanaraq ən yaxşı proqnozu seçməyə imkan verir.

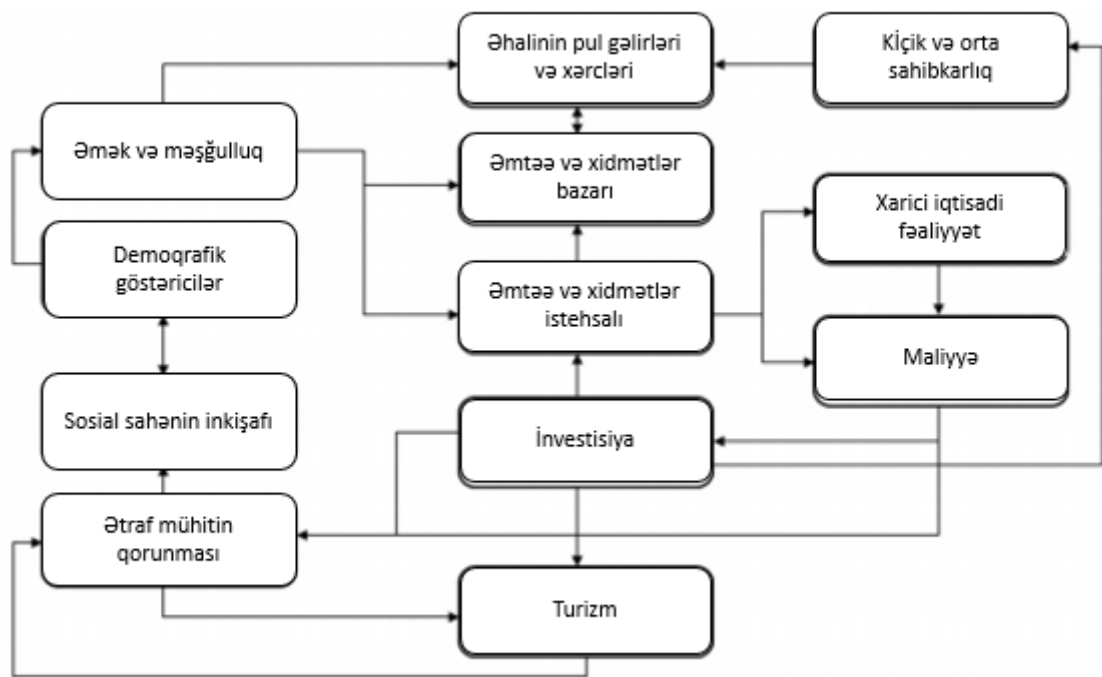
Mövcud olan müasir tendensiyalardan biri ərazi inkişafının idarəetmə sistemində, proqnozlaşdırmanın artan əhəmiyyəti və roludur. Strateji hədəf qərarları müasir proqnoz metodları və vasitələrindən istifadə etməklə əsaslandırılmalı və müəyyən bir ərazinin inkişaf xüsusiyyətlərini nəzərə almalıdır. Regional inkişafın xüsusiyyətləri nəzərə alınmadan strateji proqnozlar və planlar sadəcə boş qalacaqdır. Bölgələrin ehtiyac və xüsusiyyətlərindən kənarda heç bir uzunmüddətli milli proqram formalaşdırıla və həyata keçirilə bilməz.

Məlumatların təhlili və proqnozlaşdırılması sosial-iqtisadi proseslərin tənzimlənməsi ilə bağlı hökumət qərarlarının qəbul edilməsi üçün əsasdır. Oskar Morqanşteyn və Jon fon Neyman "Oyun nəzəriyyəsi və iqtisadi davranış" kitabında qeyd etmişdilər ki, iqtisadi nəzəriyyə bütün formalarda son nəticədə proqnoz vermək üçün hazırlanmışdır. İndiki mərhələdə dəyişiklik variantlarını daha yaxşı görməyə və məlumatlı qərarlar qəbul etməyə imkan verən etibarlı proqnozlaşdırmaya çox diqqət yetirilir. Bunun nəticəsində isə regional inkişafın idarə edilməsi elementi kimi sosial-iqtisadi proqnozlaşdırma metod və vasitələrinin öyrənilməsi aktual elmi vəzifəyə çevrilmişdir. [9]

Mərhələ 1. Tədqiqatın məqsədinin, modeldə iştirak edən izahlı dəyişənlərin məcmusunun formalaşdırılması.

Tədqiqatın əsas məqsədini dövlətin sosial-iqtisadi inkişafının göstəricilərinin proqnozunu əldə etmək təşkil edir. Dövlətin sosial-iqtisadi inkişafının göstəricilərinin ardıcıl və ziddiyyətsiz proqnozlarını əldə etmək üçün izahlı dəyişənləri seçərkən proqnoz göstəriciləri arasında qarşılıqlı əlaqələrin mövcudluğunu nəzərə almaq lazımdır. Əhali dinamikası ilə iqtisadi və sosial inkişaf parametrləri arasında sabit bir əlaqə mövcuddur. Demografik göstəricilər birincisi əmək ehtiyatlarına, ikincisi

iqtisadiyyatın inkişafına təsir göstərir və eyni zamanda əhalinin istehlak, gəlir, xərc həcminin və quruluşunun proqnozu üçün əsas rol oynayır. Sosial-iqtisadi inkişaf göstəriciləri arasındakı əlaqə Şəkil 12.-də verilmişdir. [28]



Şəkil 12. Sosial-iqtisadi inkişaf göstəriciləri arasındakı əlaqə sxemi

Mərhələ 2. Statistik məlumat.

Bu lazımi statistik məlumatları, yəni analizdə iştirak edən simulyasiya edilmiş sistemin amillərinin və göstəricilərinin dəyərlərini toplamaqdan ibarətdir. İlk məlumat aşağıdakı şərtləri ödəməlidir: [5]

- tamamlanmış olmalıdır, yəni bütün hesablama dövrü ərzində model qərarlarını formalaşdırmaq üçün kifayət qədər olmalıdır;
- modelləşdirmə məqsədlərinə çatmaq üçün kifayət qədər dəqiqlik və etibarlılığa malik olmalıdır;
- mənbə məlumatının bir hissəsi olan bütün göstəricilər bir-birinə uyğun olmalıdır, yəni bəzi məntiqi ardıcıl sistemə uyğun olmalıdır.

Retrospektivdə əsas məlumat mənbələri bunlardır:

- Dövlət Statistika Xidmətinin kolleksiyaları və hesabatları;
- Dövlət Xəzinədarlığının rəsmi hesabatları.

Hər bir göstərici üçün proqnozların hazırlanması retrospektiv dövrə görə məlumat mənbələrinin siyahısını, ekzogen amillərə görə isə retrospektiv və proqnoz dövrlərində siyahını təqdim edir.

Mərhələ 3. Seçilən proqnozlaşdırma texnikasına əsasən proqnoz göstəricilərinin hesablanması.

Sosial-iqtisadi proqnozlar hazırlanarkən aşağıdakı dörd qrup metoddan ən çox istifadə olunur:

- xətti reqressiya (əgər proqnoz göstəricisinə ən çox təsir edən amillər müəyyən edilmiş və asılılığın təbiəti xətti olarsa);
- qeyri-xətti reqressiya (proqnoz göstəricisinə ən çox təsir edən amillər və asılılığın xarakteri qeyri-xətti olduqda);
- ekstrapolyasiya üsulları (əgər retrospektiv dövrdə fərdi göstəricilərin dinamikasının təhlili göstərmişdir ki, indikatorun sonrakı dəyərləri digər amillərin dəyişməsi ilə deyil, əvvəlki dövrlərdə eyni göstəricinin dinamikası ilə müəyyən edilir);
- deterministik tənlik (hər zaman fərdi göstəricilərin dəyəri digər göstəricilərin dəyərlərinin nisbəti ilə misilsiz müəyyən edilir, buna görə də proqnoz dəyərlərini müəyyənləşdirmək üçün əvvəlcədən təyin olunmuş formullardan istifadə olunur).

Hər bir vəziyyətdə müəyyən bir tənlik növünün seçimi, retrospektiv dövr üçün göstərici dinamikasının təhlilinin nəticələrinə əsaslanır və modelin keyfiyyətini yoxlayarkən son mərhələdə də əsaslandırılmalıdır.

Mərhələ 4. Təhlil və proqnozlaşdırma üzrə hesabatların formalaşdırılması və vizuallaşdırılması.

Nümunə hasilatı və maşın öyrənməsinin ən güclü texnologiyaları belə müvafiq proqram şəklində instrumental dəstəyi olmadan heç bir əhəmiyyət kəsb etmir. Buna

görə də Deductor analitik platformasını tətbiq mühiti olaraq seçdik ona görə ki, o lazımi vasitələrə malikdir və büdcə təşkilatının fəaliyyətini təhlil etmək üçün uyğunlaşdırılmışdır. [11]

Sosial-iqtisadi inkişafın proqnoz göstəriciləri üçün məlumat anbarının konsepsiyası və arxitekturası

Deductor platformasına əsaslanan məlumatların yayılması texnologiyalarının tətbiqi dövlətin sosial-iqtisadi inkişafının proqnozunu tərtib etmə mərhələlərinin tərkibini dəyişdirir: məqsədi qeyd etmək; məqsədə uyğun bir ssenari seçmək; skript işlətmək; hesabat hazırlamaq və yükləmək; hesabatları redakte etmək. Analitikin vəzifəsi məlumatların işlənməsi ssenarisini düzgün seçimin tapılmasına qədər azaltmaq, onun icrası və hesabatını Deductor analitik platformasına verməkdir. Bu mərhələlərin həyata keçirilməsi üçün vacib bir vəzifə, müxtəlif xarici mənbələrdən, o cümlədən statistik hesabatlardan istifadə edərək vaxtaşırı doldurulan məlumat anbarı yaradılmasıdır. "Proqnozlaşdırma" məlumat anbarının tərkibinə "Göstərici", "Sənaye üzrə göstəricilər", "Təsərrüfatların kateqoriyası üzrə göstəricilər" kimi proseslər daxildir.

Deductor əlaqəli məlumat anbarının qurulması sxemi meta verilən redaktorundan istifadə etməklə həyata keçirilir. Hər bir meta verilən qovşağı üçün ölçü və ya onun atributu, yəni identifikator, etiket və məlumat növü göstərilmişdir. Sonra ölçülərin iyerarxiyası və mərkəzi fakt cədvəli ilə əlaqəsi müəyyənləşdirilir. (Şəkil 13.)

Объект	Имя
Прогнозирование	Prognostirovanie
Кубы	
Показатели	prognoz_5
Процессы	
Показатели	prognoz
Атрибуты	
Измерения	
Показатель.Код	12 pokaz_id
Год	9.0 god
Факты	
Значение показателя	9.0 znachenie
Показатели_по_отраслям	pokaz_otras
Атрибуты	
Измерения	
Показатель.Код	12 pokaz_id
Отрасль.Код	12 otras_id
Год	9.0 god
Факты	
Значение_показателя	9.0 znach_otr
Показатели по категориям хозяйств	pokaz_kateg
Атрибуты	
Измерения	
Показатель.Код	12 pokaz_id
Категория.Код	12 kateg_id
Год	9.0 god
Факты	
Значение	9.0 znachenie
Измерения	
Показатель.Код	12 pokaz_id
Атрибуты	
Показатель.Наименование	ab pokaz_name
Показатель.ЕдиницаИзмерения	ab edin_izmer

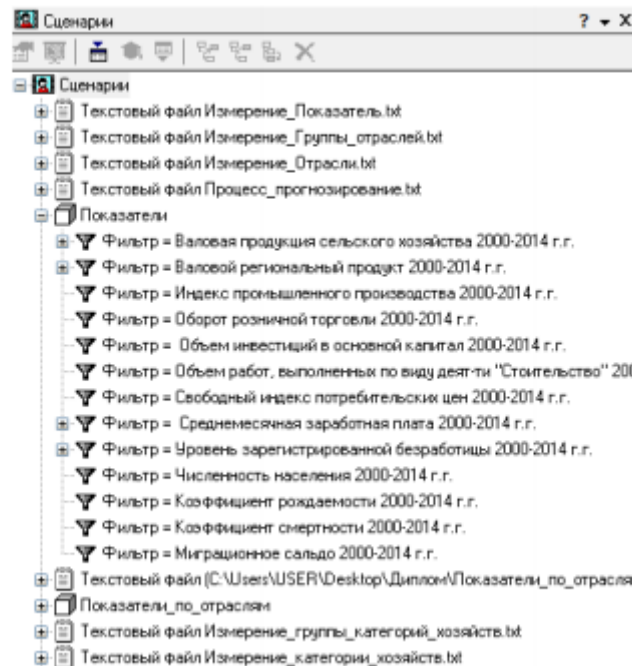
Şəkil 13. Semantik təbəqə məlumatlarının saxlanması “Proqnozlaşdırma”

Növbəti vacib addım xarici məlumatları məlumat anbarına yükləməkdir. Adi txt faylları Deductor Academic platformasının təhsil versiyasında strukturlaşdırılmış məlumat mənbəyi kimi xidmət edə bilər.

Yükləmə skriptinin ağac şəkillidir və qovşaqları aşağıdakı kimi əməliyyatlar ardıcılığıdır: [31]

- məlumatların xarici mənbələrdən idxalı;
- məlumatları ən yüksək iyerarxiya səviyyəsindən başlayaraq atributları olan ölçülərə ixrac etmək;
- “Göstərici”, “Sənaye üzrə göstəricilər”, “Təsərrüfatların kateqoriyalarına görə göstəricilər” proseslərinə məlumat ixracı.

Yükləmə skriptinin bir parçası Şəkil 14.-də verilmişdir. “Ssenari” bölməsində məlumat idxal olunur, emal, vizuallaşdırma və ya ixrac kimi funksiyalardan istifadə olunur.



Şəkil 14. Məlumat anbarı “Ssenari” paneli

Cədvəldə daxil edilmiş məlumatları təhlil etmək üçün istifadə olunan Deductor platformasının daxili analitik metodları göstərilmişdir.

Analitik metodlar

Analitik metodlar	Göstərici
Sürüşən pəncərə	Ümumi kənd təsərrüfatı məhsulu, ümumi regional məhsul
Xətti reqressiya	Ümumi kənd təsərrüfatı məhsulu
Proqnoz	Ümumi kənd təsərrüfatı məhsulu
Transformasiya (təşkilat)	Sənayə istehsalı indeksi, pərakəndə ticarət dövriyyəsi, əsas kapitala qoyulan investisiyaların həcmi, əhalinin sayı, doğum dərəcəsi, ölüm nisbəti, miqrasiya balansı
Çarpaz – cədvəl	Sənayə istehsalı indeksi, pərakəndə ticarət dövriyyəsi, əsas kapitala qoyulan investisiyaların həcmi, qeydə alınmış işsizlik səviyyəsi (iqtisadi cəhətdən fəal əhaliyə), əhalinin sayı, doğum dərəcəsi, ölüm nisbəti, miqrasiya balansı

Bütün hesablanmış proqnozlar məlumat toplusu şəklində birləşdirilir, sonra minimal səhv ilə model məlumatlarına əsasən optimal proqnoz seçilir. Beləliklə, hesablanmış məlumatlar əsasında dövlətin sosial-iqtisadi inkişafı ilə bağlı nəticələr formalaşır. Bu hesabatlar idarəetmə məlumatlarıdır və rəhbərliyi lazımi analitik və arayış materialları ilə təmin etməyə xidmət edir.

Tədqiqat zamanı müasir avtomatlaşdırılmış texnologiyalara əsaslanaraq, daxil olan məlumatların effektiv təhlili və dövlətin sosial-iqtisadi inkişafının proqnozlaşdırılması üçün sistem hazırlanır. Daxil olan məlumatların təhlili və proqnozlaşdırma prosesi təsvir olunur: statistik məlumatların toplanması, məlumat banklarının formalaşdırılması, proqnoz modelinin seçilməsi, məlumatların təhlili və işlənməsi, analitik hesabatların formalaşdırılması. Minimum səhv olan model haqqında məlumatlara əsaslanan sistem ən yaxşı proqnozu seçməyə imkan verir. Sistemdən istifadə bölgə rəhbərliyinə dəyişikliklərin variantlarını daha yaxşı görməyə və məlumatlı qərarlar qəbul etməyə imkan verən proqnoz təmin etməyə kömək edəcəkdir.

NƏTİCƏ

Sosial-ekoloji və sosial-iqtisadi sistemin tədqiqatı üçün vacib şərtlərdən biri iqtisadi göstəricilər sisteminin doğru seçilməsidir. İstifadə edilən statistik metodlar, yəni korrelyasiya, faktor təhlili və təsviri statistika davamlı inkişafın statistik tədqiqatı üçün əsas iqtisadi göstəricilər sisteminin müəyyənləşdirilməsində tətbiq oluna bilər. Statistik işlərin keyfiyyətini yüksəltmək üçün statistik göstəricilər sisteminin təkmilləşdirmək lazımdır. Seçilmiş iqtisadi göstəricilər ayrı-ayrı faktorlara görə ekonometrik modellərin qurulması üçün informasiya mənbəyi hesab oluna bilər.

Mövcud olan müasir tendensiyalardan biri ərazi inkişafının idarəetmə sistemində proqnozlaşdırmanın artan əhəmiyyəti və roludur. Məlumatların təhlili və proqnozlaşdırılması sosial-iqtisadi proseslərin tənzimlənməsi ilə bağlı hökumət qərarlarının qəbul edilməsi üçün əsasdır. Dövlətin sosial-iqtisadi inkişafının göstəricilərinin ardıcıl və ziddiyyətsiz proqnozlarını əldə etmək üçün izahlı dəyişənləri seçərkən proqnoz göstəriciləri arasında qarşılıqlı əlaqələrin mövcudluğunu nəzərə almaq lazımdır.

Daxil olan məlumatların təhlili və proqnozlaşdırılması prosesi aşağıdakı kimi təsvir olunur:

- statistik göstəricilər yığılması;
- məlumat bazalarının formalaşması;
- proqnozlaşma modelinin seçilməsi;
- analiz və emal məlumatlarının formalaşdırılması;
- analitik hesabatların formalaşdırılması.

Model məlumatlarına əsaslanan sistem minimum səhv ilə optimal proqnozu seçməyə imkan verir. Sistemdən istifadə tədqiqatçıya etibarlı proqnoz verməyə kömək edəcək dəyişiklik variantlarını daha yaxşı görməyə və məlumatlı qərarlar qəbul etməyə imkan verir.

İstifadə edilən ədəbiyyat siyahısı

Rus dilində.

1. Афанасьева, Т. В. Internet-сервис экспресс-анализа экономического состояния предприятия / Т. В. Афанасьева и др. // Двенадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2010 (20 сентября – 24 сентября, 2010 г., Тверь, Россия).

2. Афанасьева, Т. В. Применение кластеризации в электронном образовании / Т. В. Афанасьева, Н. Д. Жучков, А. Н. Солдаткин // Электронное обучение в непрерывном образовании : III Международная научно-практическая конференция : сборник научных трудов. – Ульяновск : УлГТУ, 2016.

3. Афанасьева, Т. В. Сервис прогнозирования на основе комбинирования моделей нечетких временных рядов и ARIMA / Т. В. Афанасьева и др. // Пятнадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2016 : труды конференции. В 3-х томах. Т.3 : Смоленск: Универсум, 2016.

4. Афанасьева, Т. В. Сравнение результатов регрессионной модели и модели нейронной сети при оценке стоимости нового апартаменты / Т. В. Афанасьева, К. П. Золотова, Е. А. Савенкова, М. М. Фирулина // Прикладные информационные системы : Четвертая Всероссийская НПК : сборник научных трудов. – Ульяновск : УлГТУ, 2017.

5. Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. Методы и модели анализа данных: OLAP и Data Mining. СПб.: БХВ-Петербург, 2004. 336 с.

6. Глухов, В.В. Особенности функционирования промышленности в условиях открытой экономики [Текст] / В.В. Глухов, В.Ф. Звагельский // Научно-технические ведомости СПбГПУ. Экономические науки. 2012. № 2-1 (144). С. 7—12.

7. Зинченко М. Н., Артюшина Е. А. Реализация OLAP-технологии на базе аналитической платформы Deductor Academic // Современные наукоемкие технологии. 2014. № 5 (ч. 2). С. 88-90.

8. Краснюк, Л.В. Анализ влияния промышленной политики на развитие национальной экономики [Текст] / Л.В. Краснюк // Научно-технические ведомости СПбГПУ. Экономические науки. 2012. № 6 (161). С. 16—21.

9. Morgenstern O., Дж. фон Нейман. Теория игр и экономическое поведение. М.: Книга по Требованию, 2012. 708 с.

10. Машечкин, И. В. Методы интеллектуального анализа и некоторые их приложения / И. В. Машечкин, М. И. Петровский.

11. Паклин Н. Б., Орешков В. И. Бизнес-аналитика: от данных к знаниям: учеб. пособие. СПб.: Питер, 2010. 704 с.

12. Пустыльник, П.Н. Промышленные комплексы: состояния и прогноз развития [Текст] / П.Н. Пустыльник // Научно-технические ведомости СПбГПУ. Экономические науки. 2011. № 4 (127). С. 59—64.

13. Прохоров, Е. Э. Сегментация игроков в компьютерные игры по типу поведения // Прикладные информационные системы : Третья Всероссийская НПК : сборник научных трудов / Е. Э. Прохоров ; под ред. Е. Н. Эгова. – Ульяновск : УлГТУ, 2016.

14. Регионы России. Социально-экономические показатели. 2014: Стат. сб. М. Росстат, 2014. 900 с.

15. Солонин, Е. Б. Интеллектуальные технологии поиска и анализа данных / Е.Б. Солонин // Электронное текстовое издание, УрФУ, Екатеринбург, 2015.

16. Т. В. Афанасьева, А. Н. Афанасьев Введение в проектирование систем интеллектуального анализа данных. – Ульяновск : УлГТУ, 2017.

17. Ярушкина, Н. Г. Интеллектуальный анализ временных рядов : учебное пособие / Н. Г. Ярушкина, Т. В. Афанасьева, И. Г. Перфильева. – Ульяновск : УлГТУ, 2010.

İngilis dilində.

18. Aggarwal, C. C. Data Mining: The Textbook, Springer International Publishing Switzerland, 2015.
19. Buchli, F. Detecting Software Patterns using Formal Concept Analysis, 2003.
20. Eurostat. Indicators of sustainable development. A pilot study following the methodology of the United Nations Commission on Sustainable Development. Luxembourg . Official Publications of the European Communities, 1997, Printed in Italy.
21. Eurostat. Sustainable development in the European Union. 2015 monitoring report of the EU Sustainable Development Strategy. Luxembourg, 2015
22. Eurostat. Sustainable development in the European Union. Key messages. 2013 edition.
23. Ganter, B., Wille, R. Formal Concept Analysis. In Mathematical Foundations. Springer Verlag, Berlin, 1999.
24. Gugisch, R. Many-valued Context Analysis using Descriptions. In ICCS 2001, LNAI 2120.
25. G. MARISCAL, O´. MARBA´ N AND C. FERNA´ NDEZ . A Survey of KD & DM process models and methodologies // The Knowledge Engineering Review, Vol. 25:2, & Cambridge University Press, 2010.
26. Han, J., Kamber, M. Data Mining: Concepts and Techniques // Morgan Kaufmann, 2000.
27. Joaquim L. Viegas, Susana M. Vieira, Joao M.C. Sousa. Fuzzy clustering and prediction of electricity demand based on household characteristics // International Joint Conference IFSA-EUSFLAT (16th World Congress of the International Fuzzy Systems Association (IFSA), 9th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT)), Gijon (Asturias) Spain, 2015.

28. Khanova A. A., Protalinsky O. M., Dorokhov A. F., Bolshakov A. A. Socio-Economic Systems Strategic Management Concept Based On Simulation // World Applied Sciences Journal. 2013. Vol. 24, no. 24. P. 74–79.

29. Yager, R. R. A new approach to the summarization of data // Information Sciences, vol. 28, 1982.

İstifadə edilən web saytların siyahısı.

30. <https://works.doklad.ru/view/EGSY1eobZnU.html>

31. <https://cyberleninka.ru/article/n/intellektualnyy-analiz-i-prognozirovanie-sotsialno-ekonomicheskogo-razvitiya-subekta-rossiyskoy-federatsii/viewer>

32. <https://cyberleninka.ru/article/n/analiz-ekonomicheskogo-razvitiya-i-prognozirovanie-osnovnyh-pokazateley-promyshlennosti-rossiyskoy-federatsii/viewer>

33. <http://etsim.az/upload/File/documents/02-2017/documents-16-file.pdf>

РЕЗЮМЕ

Здесь описан процесс проведения анализа поступающих данных и прогнозирования; показывает этапы процесса (формирование цели исследования, сбор участвующих объяснительных переменных, накопление важных статистических данных; анализ прогнозных показателей с использованием определенного метода прогнозирования; формирование и визуализация аналитических и прогнозных отчетов), требования к базовой информации и наиболее распространенные методы социально-экономического прогноза. Существующая взаимосвязь и взаимодействие между показателями прогнозирования должны быть приняты во внимание для получения хорошо скоординированных и согласованных данных прогнозов. На основе аналитической платформы Deductor разработана информационно-аналитическая система, целью которой является запуск сценария, предложенного аналитиком и создание отчетов. Система рассчитывает прогнозы, которые затем объединяются в один набор данных. Этот метод основан на информации о модели, которая позволяет выбрать оптимальный прогноз с минимальной ошибкой.

Summary

Here describes the process of conducting analysis of coming data and forecasting; shows the stages of the process (forming a purpose of the study, collecting participating explanatory variables, accumulation of important statistical data; analyzing forecast figures using a certain predicting method; forming and visualizing analysis and prediction reports), requirements to baseline information and most common methods of socio-economic forecast. The existing relationship and interaction between the forecasting indicators should be taken to consideration in order to obtain well-coordinated and consistent data forecasts. On the basis of the analytical platform Deductor there has been developed the information-analytical system, its purpose is running of the script proposed by an analyst, and creating reports. The system calculates the forecasts, which are then combined into one data set. This method is based on information about the model, which allows selecting the optimal forecast with a minimum error.