

ОБЩЕСТВЕННЫЙ ФОНД «ПРАВОВАЯ МИССИЯ»
МЕЖДУНАРОДНЫЙ КАЗАХСТАНСКИЙ КРИМИНОЛОГИЧЕСКИЙ
КЛУБ

Founder:
PUBLIC FOUNDATION «LEGAL MISSION»
INTERNATIONAL KAZAKHSTAN CRIMINOLOGY CLUB

ҚАЗАҚСТАННЫҢ ҒЫЛЫМЫ МЕН ӨМІРІ
НАУКА И ЖИЗНЬ КАЗАХСТАНА
SCIENCE AND LIFE OF KAZAKHSTAN

Халықаралық ғылыми-көпшілік журнал
Международный научно-популярный журнал
International popular-science journal

№3 (46) 2017

ӨНЕРТАНУ - ФИЛОЛОГИЯ
ЭКОНОМИКА - ТЕХНИКА - МЕДИЦИНА

Бас редактор — Е.О. Алауханов
«Қазақстанның еңбек сіңірген қайраткері», заң ғылымдарының докторы, профессор

Главный редактор — Алауханов Е.О.
«Заслуженный деятель РК», доктор юридических наук, профессор

Editor-in-chief — Alaukhanov Y.O.
«Honored Worker of the Republic of Kazakhstan», Doctor of Law, professor

Астана 2017

ЭКОНОМИКА ҒЫЛЫМЫ - ЭКОНОМИЧЕСКАЯ НАУКА - ECONOMICS

Бильдеббаева А.Б. ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ РАЗВИТИЯ ЭКОНОМИКИ РЕСПУБЛИКИ КАЗАХСТАН	200
Sarkulova N.K, Atenova K.A. POSITION OF KAZAKHSTAN IN THE RATING OF GLOBAL COMPETITIVENESS OF WORLD COUNTRIES	207
Bildebaeva A.B. PROBLEMS OF SMALL BUSINESS DEVELOPMENT IN THE REPUBLIC OF KAZAKHSTAN	210
Андыбаева Г.Т., Машай Н. ҰЙЫМНЫҢ ҚАРЖЫЛЫҚ ТҰРАҚТЫЛЫҒЫН БАҒАЛАУДАҒЫ ӘДІСТЕМЕЛІК ТӘСІЛДЕР	213
Китапова К., Шарымхан А. ЭВРИСТИКА ӘДІСТЕРІНІҢ ТЕОРИЯЛЫҚ МАҢЫЗДЫЛЫҒЫ	216

ТЕХНИКА ҒЫЛЫМЫ - ТЕХНИЧЕСКИЕ НАУКИ - TECHNICS

Тюленбердинова Г.А., Гусманова Ф.Р., Адилжанова С.А., Газиз Г.Г., Найманбай А.А., Алтыбай А. DATA MINING-ТІҢ НЕГІЗГІ МӘСЕЛЕЛЕРІ.....	221
Бегимбетова Х.А. SCADA БАҒДАРЛАМАСЫНЫҢ БАҒДАРЛАНҒАН-ҚҰРАУШЫ СӘУЛЕТІ	229
Жаптаев А.А., Назарбек Т.С. ФИШИНГ ШАБУҒЫ	236
Назарбек Т.С., Гайруллаева Р.М. ТРИГОНОМЕТРИЯЛЫҚ ТЕНДЕУЛЕР ЖҮЙЕСІН ШЕШУ ӘДІСТЕРІ	241

МЕДИЦИНА ҒЫЛЫМЫ - МЕДИЦИНСКАЯ НАУКА - MEDICAL SCIENCE

Нугманова А.М., Ашимова М. Г., Базилова А.А., Джахудаева А.Н., Ниязбаева А.С., Джахудаева А.Н., Рознева С.К. ОСТРОЕ ПОЧЕЧНОЕ ПОВРЕЖДЕНИЕ У ДЕТЕЙ	246
Тюленбергенова Б., Анарбаева А.У АНА СҮТІНІҢ ЖАҢА ТУҒЫЛҒАН СӘБИДІҢ ДЕНСАУЛЫҒЫ ҮШІН МАҢЫЗДЫЛЫҒЫ	249



**ТЕХНИКА ҒЫЛЫМЫ
ТЕХНИЧЕСКИЕ НАУКИ
TECHNICS**



Төлебердинова Гульнур Аллысқызы

Әл-Фараби атындағы ҚазҰУ, механика-математика факультеті,
информатика кафедрасының доценті, ф.-м.ғ.к

Гусманова Фаридә Равиловна

Әл-Фараби атындағы ҚазҰУ, механика-математика факультеті,
информатика кафедрасының доценті, ф.-м.ғ.к

Адижанова Салтанат Альмуханбетқызы

Әл-Фараби атындағы ҚазҰУ, механика-математика факультеті, информатика кафедрасының
аға оқытушысы

Газиз Гүлнұр Газизқызы

Әл-Фараби атындағы ҚазҰУ, механика-математика факультеті,
информатика кафедрасының аға оқытушысы

Найманбай Асан Аманжолұлы

Әл-Фараби атындағы ҚазҰУ, механика-математика факультеті, дифференциалдық теңдеулер
және басқару теориясы кафедрасының ассистенті

Алтыбай Аршын

Әл-Фараби атындағы ҚазҰУ, механика-математика факультеті, дифференциалдық теңдеулер
және басқару теориясы кафедрасының оқытушысы, a_arshyn@mail.ru

DATA MINING-ТІҢ НЕГІЗГІ МӘСЕЛЕЛЕРІ

Түйін. Бұл мақалада деректерді интеллектуалды талдаудың негізгі мәселелерінің бірі болып табылатын жіктеу алгоритмдері мен жіктеу есебін шешу сұлбасы қарастырылады. Жіктеу есебіндегі шешім ағашын пайдаланатын алгоритмдер жан-жақты көрсетіледі.

Кілттік сөздер: DATA MINING. Жіктеу. Деректер. Талдау

Summary. This article deals with the classification and algorithms using decision trees which are the main tasks of DATA MINING

Keywords: DATA MINING, classification. Data. Analysis

Резюме. В этой статье рассматриваются классификация и алгоритмы, использующие деревья решений которые являются основными задачами DATA MINING

Ключевые слова: DATA MINING; Классификация; Алгоритм; Данные. Анализ

Сандық ақпараттарды жинаудың заманауи әдістері мен оны сақтау технологиясының дамуына үлкен ДҚ-ның тез таралуына әкелді. Бұл адамзат қызметінің барлық салаларында, дәстүрлі салалардан (мысалы, кредиттік карточкамен сатып алулар туралы жазбалар, телефон қоңыраулары туралы

лық құжаттамалар) салаларға дейін пайда болды. ДҚ-н пайдаланушылар үшін маңызды мәні болуы мүмкін деректерді жүйелендіру мүмкіндіктерін оқуға және одан ақпарат алуға қызығушылығы арта түсті. Осындай зерттеулермен байланысты сала (ағылш. Data Mining – деректерді интеллектуалды талдау, DM) атымен белгілі болды.

DM үлкен ДҚ-нан болжамдық ақпараттарды генерациялау және автоматты түрде шығару үрдісін орындауды білдіреді. Ол пайдаланушыларға бір уақытта пайдалы және түсінікті болатын, бұрын белгісіз ара қатынастағы айқын емес деректерді табу мақсатында бақылаулар, деректер жиынын әдістерін талдауды қамтиды. Бұл мақалада біз деректерді интеллектуалды талдаудың негізгі мәселелері қарастырылады.

Жіктеу

Деректердің интеллектуалды талдауының ең белгілі әдісі – жіктеу болып табылады. Қосымшалардың көптігімен сипатталады: суреттер мен шаблондарды анықтау, медициналық диагностика, кредиттерді мақұлдау мен бағалау, қылмыстың сұлбаларын анықтау, қаржылық тенденцияларды шығару және т.б. Көптеген зерттеушілердің айтуынша болжау мен бағалауды, жіктеудің түрлері

1-кесте. Білімді бағалаудың негізгі шкаласы мен критерийлері

Жинаған баллдың %-ы (100-ден)	Бағаның тапбасы	Дәстүрлі бағалау
95-100	A	Өте жақсы (Excellent)
90-94	A-	
85-89	B+	Жақсы (Good)
80-84	B	
75-79	B-	Қанағаттанарлық (Satisfactorily)
70-74	C+	
65-69	C	
60-64	C-	
55-59	D+	
50-54	D	Қанағаттандырарлықсыз (Unsatisfactorily)
0-49	F	

1-кесте оқытушылардың оқушыларды бағаларына сәйкес қалай жіктейтінін түсіндіреді. Бұл жіктеулерді оқушының жинаған баллына шек қою арқылы алуға болады. Сондықтан, жіктеу өз кезегінде алдын ала анықталған топтар және кластармен шамаларды қою арқылы деректерді бөлетін әдіс пен үрдістен тұрады деп айтуға болады.

1-анықтама. $D = \{t_1, t_2, \dots, t_n\}$ ДҚ мен $C = \{C_1, C_2, \dots, C_m\}$ кластар жиыны берілсін. Жіктеу есебі $f: D \rightarrow C$ бейнелеуін анықтауды талап етеді, мұндағы қордың әрбір элементі кластардың біріне тиесілі. Әрбір C_j класы тек соған сәйкес қортеждерін ғана қамтиды, яғни:

$$C_j = \{t_i | f(t_i)\}, 1 \leq i \leq n, t_i \in D.$$

Жіктеу алгоритмдері жіктеу үшін тек деректердің атрибуттарының мәні және жіктеу үрдісі екі фазаға бөлінеді деп жорамалданады:

1. үйретуші деректерді бағалау үшін модельді әзірлеу;
2. зерттелетін ДҚ-нан қортеждер (жолдар) жіктелуі.

Ескерту: Жаттықтыратын деректер кіріс деректері мен оларға сәйкес кластардан тұрады.

1-ші суретте жіктеу үрдісін үйлестіреді: а) сегмент кеңістігін бөлу жолымен алдын ала анықталған кластарды суреттейді, ә) деректердің кіріс жиынын қамтамасыз етеді, б) анықталған кластар негізінде деректердің жіктелуін көрсетеді (яғни, әрбір объект – нүкте белгілі бір класқа тиісті).



If ($X_4 = 1$) THEN I Класс ; If ($X_4 = 1$) THEN II Класс ; If ($X_4 = 0$)
THEN I Класс;
If ($X_4 = 0$) THEN II Класс;

Осыдан кейін, бұл ережелердің 3-кестеге сәйкестігін тексеру керек. Онда кателіктердің көп мөлшерде екені белгілі, яғни кестеде көрсетілген класс ережесі бойынша есептелген нәтижелердің сәйкессіздігі. Мысалы, I ереженің нәтижесі ($X_4 = 1$ предикатымен) 5, 7, 8 қортеждеріндегі берілген класпен сай келмейді. Барлық мүмкін болатын I предикаты бар (X_1, X_2, \dots, X_8 үшін) ережелер тексеріліп болған соң, барлық қортежді дұрыс жіктеуге бұл ережелердің жеткіліксіз екендігі айқындалады. Келесі қадам барлық екі предикатты мүмкін болатын

IF ($X_i = K$) \cap ($X_j = M$) THEN N Класс;

ережелерді пайдаланумен байланысты, мұндағы (i, j) барлық мүмкін болатын мәндерді қабылдайды ($1, \dots, 8, (i \neq j)$), ал K, M – 0-ге немесе 1-ге тең. Тап осындай тексеруден соң жіктеуді дұрыс орындайтын ережелер қатарын табамыз. Соңғы қадам осындай ережелер жиынынан оңтайлы (минималды) шешімді таңдауға негізделеді. Әрине, бұл сипаттамалар тек Жабу алгоритмі идеясының түсіндірмесі ғана. Сондай-ақ, нейрондық желілерді пайдаланатын логикалық ережелердің генерациялық алгоритмдері де кездеседі.

Қорытынды

Біз бұл мақалада деректерді интеллектуалды талдаудың негізгі мәселелерінің бірі болған жіктеу алгоритмдерін нақты мысалдар арқылы сипаттадық және жіктеу есебін шешу сұлбасы графикалық бейнелеу элементтері арқылы толығымен қарастырдық. Жіктеу есебіндегі шешім ағашын пайдаланатын алгоритмдерді, анықтамалары, графикалық бейнелеу элементтері және нақты мысал келтіру арқылы толығымен қарастырдық.

Әдебиеттер тізімі

1. *Emerging cubes for trends analysis in OLAP databases / S. Nedjar [et al.] // DaWaK. Lecture Notes in Computer Science. Springer, 2007. Vol. 4654. P. 135–144. DOI: 10.1007/978-3-540-74553-2_13.*
2. В.Дюк, А. Самоіленко. *Data mining- учебный курс.*, 2001, Сп.Петербург, 367с.
3. Margaret H.Dunham. *Data mining. Introductory and Advanced Topics. Pearson Education, Singapore 2003; Taj press, India 2005.* , 314pp.
4. *Методы и модели анализа данных: OLAP и Data Mining / А. А. Барсегян [и др.]. СПб.: БХВ-Петербург, 2014. 336 с.*
5. Миронов В. В., Юсупова Н. И. *Концептуальные модели баз данных. Многомерные модели.* Уфа: УГАТУ, 2010. 83 с
6. *Microsoft SQL Server 2005 Analysis Services. OLAP и многомерный анализ данных / А. Б. Бергер [и др.]. СПб.: БХВ-Петербург, 2011. 928 с.*
7. Harinath S., Quinn S. *Professional SQL Server Analysis Services 2005 with MDX.* N. Y.: Wiley, 2007. 848 p.