

Маматов Нарзулла Солиджанович

Т.ф.д., ТАТУ ҳузуридаги ахборот – коммуникация технологиялари илмий – финновацион маркази катта илмий ходими

Эл. почта: m_narzullo@mail.ru

Бобоев Лочинбек Боймуротович

ТАТУ АТДТ кафедраси ассистенти

Эл. почта: repytolochin@gmail.com

Б.И.Отахонова

ФарДУ, АТ кафедраси ўқитувчиси

Эл. почта: bahrixon@mail.ru

Babomuradov O.J., Mamatov N.S., Boboev L.B, B.I.Otahonava

Text preprocessing and classification problem on the basis of probabilistic estimation models

This paper analysis results of theoretical and practical research of classification of textual data presented in various natural languages. Approaches to research have been determined based on the analysis of the sources. Approaches to the formation of informative features by preprocessing of the text are given for the classification of textual electronic documents based on probability models. The model of the textual documentation was evaluated by Bernoulli effectiveness and based on the results of the experimental research.

Keywords: text documents, preprocessing text data, classification, Bayes probability, Bernoulli distribution, analysis, feature space, vocabulary, train data, words.

Эл. почта: repytolochin@gmail.com

УДК 621.396.41

М.Б. Зайнутдинова, Ш.Б. Сайфуллаев, Н.Ф. Ахмедова

КАТТА ҲАЖМЛИ МАЪЛУМОТЛАРНИ ҚАЙТА ИШЛОВЧИ АХБОРОТ ТИЗИМИ ЁРДАМИДА КЎКРАК БЕЗИ САРАТОНИНИ АНИҚЛАШ

Мақолада кўкрак беzi саратонини хавfli риск гуруҳини шаклланишини ва индивидуал риск факторларини аниқлаш ҳамда таҳлил бўйича ахборот тизимини ишлаб чиқиш натижалари келтирилган. Кўкрак беzi саратони rischi мавжуд энг аҳамиятли факторларнинг ўрнатиш имконини берувчи илк натижалар олинган.

Калит сўзлар: кўкрак беzi саратони, ахборот тизими, хавfli риск гуруҳи, риск факторлари, башорат жадвали, анкета статистика, онкология.

Онкология соҳасининг муаммоларидан бири Ўзбекистон аёлларининг кўкрак беzi саратони (КБС) касаллиги билан касалланган орасида юкори позиция эгаллаб турувчи кўкрак беzi саратони касаллигини профилактикаси ҳисобланади [6]. Шу сабабли бу касалликни профилактика режасида касалликнинг ривожланишига таъсир қилувчи факторлар комплексини аниқлаш асосида КБС хавfli риск ва индивидуал риск гуруҳларини шаклланишини баҳолаш билан “соғлом” аёллар контингенти орасида касалликнинг эрта

боскичларида кўкрак беzини ҳосиласини аниқлаш долзарб ҳисобланади. Этиологик факторларнинг аниқлаш усулларини бири соғлиққа зарар етказмасдан туриб аҳоли контингентини кенг камраб олувчи, ҳудудий хусусиятларни ҳисобга оладиган, амалга оширишда содда бўлган чуқур изланиш учун бирламчи танлов ўтказиш имконини бурувчи анкеталаштириш усули ҳисобланади. Катта ҳажмдаги анкета сўрови маълумотлар массивини қайта ишлаш учун компьютер технологияларидан фойдаланиш истиқболли ҳисобланади.

1- жадвал

Республика вилоятлари бўйича КБС билан касалланиш даражаси

Вилоятлар	100 000 аҳолига интен-сив кўрсат-кич	100 000 аҳолига (дунё) стандартлаштирилган кўрсаткич		
		Умумий	Шахар	Қишлоқ
		M±m	M±m	M±m
Қорақалпоғистон Республикаси	5,6	8,2±1,7	15,2±5.7	5,6±1.1
Андижон	7,7	9,5±1,3	12,8±2.1	10,1±2.3
Бухоро	8,1	11±1,2	13,9±3.4	10,6±0.3
Жиззах	6,3	10,2±1,7	7±3.1	9,6±2.4
Қашқадарё	5,7	7,5±0,6	9,2±0.6	7,5±0.7
Навоий	7,4	12,3±4,2	9±3.8	12,2±1.6
Наманган	6,0	8,2±1,7	7,8±2.8	9,3±1.6
Самарқанд	6,3	8,1±1,0	14,1±1.1	7,2±0.8
Сирдарё	6,6	12,2±5,6	14,7±3.1	7,5±2,8

Сурхондарё	4,2	6,3±1,4	4,7±1.6	5,9±0,6
Тошкент	9,0	11±0,7	19,9±9.5	10,2±.5
Фарғона	7,4	9,2±0,9	22,8±7.1	6,3±2.3
Хоразм	6,4	9,3±0,7	9,6±0.9	8,5±0.9
Тошкент шаҳри	19,5	22,5±2,9	25,2±0.9	-
Республика бўйича жами	7,6	10,4±1,2	13,2±6.0	8,5±2.0

Ўзбекистон Республикаси Президенти Шавкат Мирзиёев ташаббуси билан 2017-йил 6-июлида Тошкент шаҳрида соғлиқни сақлаш тизимини ривожлантириш, мавжуд муаммоларни ва соҳадаги камчиликларни бартараф этишга бағишланган видеоселектр йиғилиши ўтказилди. Давлат раҳбари имзо чеккан 2017-йил 4-апрелдаги ПҚ-2866-сон «2017-2021 йилларда Ўзбекистон Республикасида онкология хизматини янада ривожлантириш ва аҳолига онкологик ёрдам кўрсатишни такомиллаштириш чора-тадбирлари тўғрисида»ги қарори аҳоли орасида онкологик касалликларни эрта аниқлаш ва олдини олишга йўналтирилган.

Тадқиқотнинг мақсади КБС хавфли риск гуруҳини шакллантириш ва индивидуал риск факторларини аниқлаш ҳамда таҳлил бўйича анкеталаштириш натижаларини умумлаштириш учун ахборот тизимини (АТ) ишлаб чиқиш ҳисобланади.

Тадқиқот усуллари ва материаллари

Тадқиқот ишининг асоси бўлиб турли ҳудудларда яшовчи аёллар орасида КБС ни ривожланиш rischi факторларини аниқлаш бўйича анкетали сўров натижаларининг таҳлили ётади. Тадқиқот КБС билан кассалланган ва соғлом аёллар орасида “ҳолат-назорат” усули бўйича ўтказилди. Республикадаги 2015-йилдаги КБС касаллиги билан касалланиш ҳақидаги маълумотлар Ўзбекистон Республикаси Давлат комитети сайтидан олинди [7].

Юқоридаги маълумотлар манбаи мавжудлигини ҳисобга олган ҳолда Россия онкология марказларида жорий қилинган дастурий платформаларни асос қилган ҳолда Ўзбекистон шароити учун ҳам махсус платформа ишлаб чиқишда Россия тажрибасини ўрганиш ва шу асосда платформа ишлаб чиқиш маъқбул деб ҳисобланади. Қуйида Россияда жорий этилган тизимнинг яратилиши ва жорий қилиниши ҳамда дастлабки натижалар таҳлил қилинди.

Алоҳида тадқиқотларда олинган риск факторларини умумлаштириш ва аниқлаш учун ҳамда КБС касаллигини ривожланиш riskлари билан боғлиқ умумлашган факторлари сабабли чораларни аниқлаштириш учун таҳлилланувчи натижаларни статистик аҳамиятини ошириш имконини берувчи мета-таҳлил усули қўлланилинган [2]. Олинган маълумотларнинг катта қисми сифатли (номинал ва тартибли) бўлганлиги сабабли анкеталаштириш натижаларини қайта ишлаш учун Woolf усули қўлланилди [3, 4, 5]. Мета-таҳлил натижалари бўйича сифатли аломатлар учун Байес классификатори асосида КБС риск гуруҳини шакллантириш учун башорат қилувчи модел қурилган [1]. Аломатлар сифатида анкета саволлари

билан аниқланувчи риск факторларидан фойдаланилинган. Моделга касаллик келиб чиқиш rischi билан алоқадор статистик аҳамиятли факторлар киритилинган. Касалланган ва соғлом гуруҳларга ажратишнинг оптимал чегараси олиш учун операцион характерли бурилиш (ROC) қурилган. Бошорат қилувчи моделнинг ташхис самарадорлигини баҳолаш сезувчанлик, хоссавийлик, ташхисий аниқлик, ижобий ва салбий натижанинг башорат аҳамияти кўрсаткичлари бўйича амалга оширилган.

Республика вилоятлари бўйича КБС билан касалланиш даражаси Натижалар

АТ ишлаб чиқиш учун КБС ривожланиш риск факторларини аниқлаш бўйича анкета таркиби ва тузилиши таҳлил қилинди. Анкета саволларини тузиш эндоген ва экзоген риск факторлар гуруҳларини аниқлашга йўналтирилган. Ҳудудий хусусиятларга (аҳоли зичлиги даражаси, миграция оқимининг жадаллиги ва бошқалар) боғлиқ ҳолда турли тадқиқотларда фойдаланилган анкеталар ўзаро тузилиши, сони (95 дан 120 тагача), саволлар тоифаси ва аломатлар градацияси билан бири-бирдан фарқ қилади.

Анкета тузилишини таҳлил қилган ҳолда АТ ишлаб чиқишга қўйиладиган талаблар шакллантирилган:

- Яратилувчи анкета тузилиши ва саволлар шакллантирилган тавсифи
- Анкеталаштириш натижаларига биноан маълумотлар базасини тўлдириш;
- Ҳам алоҳида, ҳам турли тадқиқотлар жамланмаси натижалари таҳлили асосида риск факторларини аниқлаш;
- Хавфли риск гуруҳларини шакллантириш.

Ахборот тизимини шакллантириш асоси бўлиб, қоғоз кўринишидаги анкеталар хизмат қилди. Тизимда учта база қўзда тутилган: саволлар ва жавоблар тавсифининг базаси, анкета бланкалари базаси ва анкеталаштириш натижаларининг базаси. Саволлар ва жавоблар тавсифининг ягона базаси турли тоифадаги анкеталар бланкасини шакллантириш учун керакли барча ахборотни ўзида сақлайди. Анкеталар бланкалари базасида анкета тузилиш бўйича ахборот ва ҳар бир тадқиқот номи сақланади. Анкета бланкаси бўйича анкеталаштириш базасига қоғоз бланкасидаги ёки бошқа ахборот ташувчидаги маълумотлар киритилади.

Риск факторларини аниқлаш (ҳам битта, ҳам бир нечта тадқиқот учун) учун тадқиқотнинг натижалари базасига юборилган сўровлар амалга

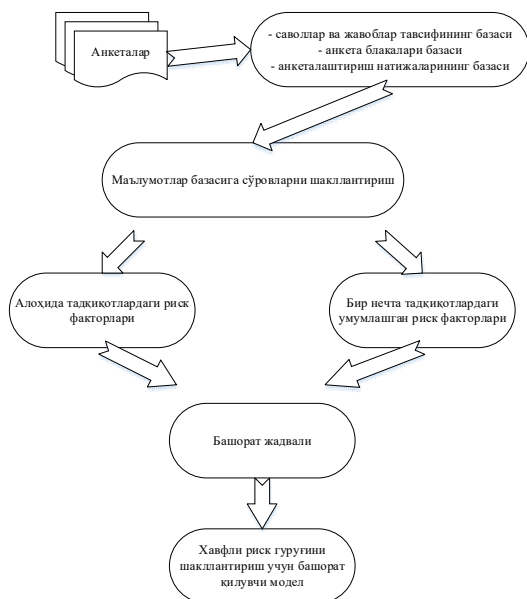
ошириш йўли билан бирламчи маълумотлар танлови шакллантирилади. Риск факторларини аниқлаш учун маълумотлар таҳлилланади, бу бўйича башорат жадвали курилади ва хавфли риск гуруҳи шакллантирилиши ўтказилади.

Анкеталаштириш натижаларини таҳлили асосида КБС касаллигининг ривожланиш риск факторларини аниқлаш бўйича АТ нинг тузилмавий схемаси 1-расмда келтирилган.

Тизимни ишлаб чиқиш натижасида анкеталар таркиби бўйича турли маълумотларни ягона соддалаштирилган базага тўплаш ва сақлашни, КБС касаллигини ривожланишининг умумлашган риск факторларини аниқлаш бўйича таҳлилни ўтказишни самарали ташкиллаштириш имкониятига эришилади

Ахборот тизими иккита дастурий модул кўринишида амалга оширилган: “ПИФАРО” ва “РИСК”.

“ПИФАРО” модулининг имкониятлари куйидагилардан иборат: анкета тузилишини ишлаб чиқиш; қоғоз анкетадан маълумотларни киритиш учун анкетанинг экран шаклини яратиш; компьютерга маълумотлар билан тўлдирилган қоғоз анкета маълумотларини киритиш, маълумотларни статистик қайта ишлаш (мета-таҳлил усули), киритилган анкеталар ва ҳисоб-китоб натижалари экспорти. “ПИФАРО” АТ модули учта асосий блоклардан иборат: “Тахрирлаш”, “Анкеталаштириш” ва “Статситика”.



1- расм. Ахборот тизимининг тузилмавий схемаси

КБС касаллигини умумлашган ривожланиш риск факторларига куйидагилар таалуқли деб белгиланган: ёш (50-59 ва 60-80 ёшли аёлларда касалликни ривожланиш rischi мос ҳолда 2,3 (1,7-3,1) ($rtet=0,3$) ва 9,2 (6,1-13,8) ($rtet=0,7$) ни ташкил этади); миллат (рус аёллари бошқа миллат аёлларига нисбатан 3,2 марта ($rtet=0,4$) кўп касланган, тувин, корейс аёлларининг кўрсаткичи анча паст ($ORsum=0,3$ (0,2-0,4); $rtet= -0,3$)); насл (онаси ёки

опа-сингиллари КБС касалланган аёллар орасида касаллик ривожланиш rischi мос ҳолда 4,2 ($rtet=0,5$) ва 5,1 ($rtet=0,6$) марта ошган); касаллик келиб чиқиши билан боғлиқ эндокрин-метоболик факторлар (анамнез фиброаденоматоз борлигида КБС ривожланиш rischi 2,4 мартага ($rtet=0,3$), сут беши жароҳатида 6,6 мартага ($rtet=0,6$), гипертоник касалликларда 2,2 мартага ($rtet=0,3$) ошган, оғирлиги 60 кг гача бўлган аёлларда КБС пасайган ($ORsum=0,5$ (0,4-0,7); $rtet= -0,3$), оғирлиги 80 кг дан юқори бўлган аёлларда 3,1 марта ошган ($rtet=0,4$)); репродуктив анамнез хусусиятлари (доимий менструацияда КБС rischi пасайган ($ORsum=0,3$ (0,2-0,5), $rtet= -0,4$); менструация йўқлигида ($ORsum= 3,4$ (2,4-4,8); $rtet=0,5$), уларнинг доимиймаслигида ($ORsum= 1,5$ (1,0-2,5); $rtet=0,2$), кечки биринчи туғруқ – 26-30 ёшдан юқори ёшда ($ORsum=1,8$ (1,1-3,2); $rtet=0,2$), кечки монопауза – 50 ёшда ва ундан юқори ёшда ($ORsum=3,0$ (1,9-4,6); $rtet=0,4$), менопауза давомийлиги 5 йилдан юқори бўлганда ($ORsum=3,1$ (2,4-3,9); $rtet=0,4$) КБС ривожланиш rischi ошади); овқатланишнинг ўзига хослиги; ижтимоий-иқтисодий статус; миграцион кўчиш; касбий зарарни мавжудлиги.

Аниқланган умумлашган факторлар КБС касаллигини хавфли риск гуруҳини шакллантириш ва респондент индивидуал riskини баҳолаш учун мўлжалланган “РИСК” ахборот тизимини модулини ишлаб чиқишда фойдаланилган КБС ривожланиш индивидуал riskини баҳолаш бўйича башорат моделини куришга асос бўлди

Шундай қилиб, таклиф қилинаётган КБС хавфли риск гуруҳини шакллантириш башорат модели ва ахборот тизимини ишлаб чиқиш хавфли риск гуруҳини шакллантириш, автоматлаштирилган синовдан ўтказиш натижалари бўйича аёлларни соғлом ва касалланган гуруҳларга самарали ажратиш имконини беради. Бу чуқур тадқиқот ўтказишга оид шахслар доирасини торайтириш имкониятини бериш, ўз вагида тадқиқот ўтказиш ва аёлларни даволашни назарда тутати, шифокорни вақтини ва ташхис учун керакли восталарни тежайди. Дастурдан фойдаланиш КБС касаллигини эрта аниқлаш имкониятини беради, натижада ўлим кўрсаткичи пасаяди, башорат яхшиланади ва аёлларнинг яшаш сифати ошади.

Ишлаб чиқилган КБС риск гуруҳини шакллантиришни башорат модели ва ахборот тизими исталган ҳудудда, айниқса аҳоли пунктлари марказидан узоқда яшовчи аҳолини оммавий скрининг қилишнинг биринчи босқичи сифатида фойдаланиш учун тавсия этилади.

Адабиётлар:

1. Гублер Е.В. Вычислительные методы анализа и распознавания патологических процессов. Л.: Медицина, 2015. 296 с.

2. Плавинский С.Л. Руководство по мета-анализу: учебное пособие. СПб.: СПбМАПО, 2014. 75 с.

3. Сергиенко В.И., Бондарева И.Б. Математическая статистика в клинических исследованиях. М.: ГЭОТАР-МЕД, 2015. 256 с.

4. Аксель, Е.М. Злокачественные новообразования молочной-железы: состояние онкологическое помощи; заболеваемость и смертность / Е.М. Аксель/ /Маммология. - 2006.

5. Ананина, О.А. Ранней диагностика рака молочной железы на основе информационной системы. / Диссертация/ - 2008

6. Очик маълумотлар портали – <http://www.data.gov.uz>

7. Ўзбекистон Республикаси Давлат комитети сайти – <http://www.stat.uz>

Зайнутдинова Мастура Баходировна

т.ф.н., Муҳаммад ал-Хоразмий номидаги Тошкент ахборот технологиялари университети (ТАТУ) Ахборот технологиялари (АТ) кафедраси доценти.

Сайфуллаев Шерзод Бахтиёр ўғли

Муҳаммад ал-Хоразмий номидаги ТАТУ АТ кафедраси магистранти

Эл. почта: shezodsay@gmail.com

Ахмедова Нозима Фарход кизи

Муҳаммад ал-Хоразмий номидаги ТАТУ АТ кафедраси магистранти

Эл. почта: Rachel_051294@mail.ru

Zaynutdinova M.B., Sayfullaev Sh. B., Akhmedova N.F.

Detection of breast cancer using information system which process big data

The article presents the results of the development of an information system for the analysis and identification of individual risk factors and the formation of groups of increased risk of breast cancer. The first results were obtained, which allowed to establish the most significant risk factors for breast cancer.

Keywords: *breast cancer, information system, hazardous risk group, risk factors, forecasting table, questionnaire statistics, oncology.*

М.Б. Зайнутдинова, Ш.Б. Сайфуллаев, Н.Ф. Ахмедова

Обнаружение рака молочной железы с использованием информационной системы, которая обрабатывает большие данные

В статье представлены результаты разработки информационной системы по анализу и выявлению факторов индивидуального риска и формированию групп повышенного риска рака молочной железы. Получены первые результаты, позволившие установить наиболее значимые факторы риска рака молочной железы.

Ключевые слова: *рак молочной железы, информационная система, группа риска, факторы риска, таблица прогнозирования, статистика вопросников, онкология.*

УДК 004.056.53

Рахманов А.Т., Керимов К.Ф., Камалов Ш.К.

АЛГОРИТМ АВТОМАТИЧЕСКОГО ОБНАРУЖЕНИЯ УЯЗВИМОСТИ ВИДА SQL ИНЪЕКЦИИ

В данной статье разрабатывается алгоритм обнаружения атаки по инъекции SQL с помощью функции извлечения одного символа, и оценку эффективности предложенного алгоритма с помощью искусственных данных.

Ключевые слова: SQL инъекция, уязвимость, обнаружение атак, определение угроз, использование специальных символов и строк, алгоритм.

Математическое моделирование и идентификация информационных объектов играют важную роль при решении задач распознавания образов. Одной из таких задач является обнаружение атак или нормальных запросов на веб приложения. Исследования, посвящённые изучению обнаружения атак или нормальных запросов на веб приложения, начались сравнительно

недавно. Но, тем не менее, существует много исследований в этом направлении [1-11]. Применение математических методов в решении таких задач в основном велись японскими учеными [1,2,6]. Например, в работе [1] предложено 2 способа обнаружения атак SQL инъекций, основанных на свойстве распределения символов при построении атак SQL инъекций. В работе [2]