ТУРЛИ ТИПДАГИ МАЪЛУМОТЛАР ФАЗОСИДА ФИШЕР ТИПИДАГИ МЕЗОНГА АСОСЛАНГАН ИНФОРМАТИВ БЕЛГИЛАР МАЖМУАСИНИ ТАНЛАШ АЛГОРИТМИ

Нишанов Ахрам Хасанович[1], Актамов Шоҳрухбек Улуғбек ўғли[1], Олламберганов Файзулла Фарход ўғли[1]

[1]Муҳаммад-ал Хоразмий номидаги Тошкент ахборот технологиялари университети, Ўзбекистон, Тошкент
e-mail: nishanov_akram@mail.ru, fayzulla0804@gmail.com

K E Y W O R D S

Фишер мезони, информатив белгилар танлаш, турли типдаги маълумотлар, объектлар ўхшашлиги, оптимизация масаласи, эвристик функционал, номинал ва қийматли белгилар.

A B S T R A C T

Мақолада Фишер типидаги мезонга асосланган информатив белгиларни танлаш алгоритми турли типдаги маълумотлар фазосидаги объектларнинг ўхшашлик даражасини аниқлаш ва улардан оптимал белгилар мажмуасини танлашга йўналтирилган. Алгоритмнинг асосий мақсади – берилаётган маълумотлар фазосидаги номинал ва қийматли белгилар асосида объектларнинг ўзаро ўхшашлик ва фарқларини ҳисоблаб, улардан энг муҳимларини аниқлаш. Бунинг учун алгоритм Фишер типидаги мезондан фойдаланади, у синф объектларининг ўртача ўхшашлик даражасини ҳисоблашга ёрдам беради. Алгоритм босқичма-босқич ишлаб, белгиланган мақсад функциясини оптималлаштириш орқали информатив белгилар мажмуасини топишга қаратилган. Шунингдек, алгоритм турли типдаги маълумотларни ҳисобга олиш имкониятига эга бўлиб, бу уни кўпқиррали ва самарали қилади. Эвристик мезонлар ёрдамида алгоритм синф объектларининг ўзаро ўхшашлик даражасини баҳолашда аниқ ва самарали натижаларга эришади. Бунда, оптимизация масаласи қийматларининг кичиклиги синфлар орасидаги фарқларни аниқлашда муҳим аҳамият касб этади. Ушбу алгоритмнинг асосий афзаллиги шундаки, у турли типдаги маълумотлар фазосида самарали белгиларни танлаш имконини беради, бу эса таҳлил жараёнини яхшилайди ва маълумотлардан самарали фойдаланишга ёрдам беради.

Кириш

Фишер мезонига асосланган информатив белгиларни танлаш алгоритмлари замонавий маълумотлар таҳлили соҳасида муҳим ўрин тутади. Ушбу алгоритмлар турли типдаги маълумотлар фазосида оптимал белгиларни аниқлаш ва уларни классификация ёки бошқа таҳлил жараёнларида қўллашда самарали ҳисобланади. Информатив белгиларни аниқлаш жараёни кўплаб соҳаларда, жумладан, биомедицина, иқтисодиёт ва муҳандисликда қўлланилиб, юқори натижадорликка эришиш имкониятини яратади [1-7].

Информатив белгиларни танлаш жараёнида Фишер мезонининг афзалликларидан бири бу – объектлар ўртасидаги ўхшашлик ва фарқларни ҳисоблаш имкониятидир. Бу мезон орқали белгиларнинг аҳамият даражаси аниқланади ва таҳлил учун энг муҳим бўлганлари танлаб олинади [3-9]. Бу жараён маълумотлар фазосидаги турли типдаги объектлар учун самарали белгилар мажмуасини танлаш имконини беради.

Алгоритмнинг асосий босқичларидан бири бу – синф объектларининг ўзаро ўхшашлик даражасини аниқлаш ва улардан оптимал белгиларни танлашдир [18-24]. Бу босқичда эвристик мезонлар қўлланилади, улар синф объектларининг ўзаро муносабатларини аниқлашда муҳим ўрин тутади [4,10-17]. Бу мезонларнинг қийматлари синфларнинг ўзаро ўхшашлик ёки фарқларини аниқлаш имконини беради.

© *А.Х. Нишанов, Ш.У. Актамов, О.Ф. Олламберганов*

Шу сабабли, ушбу алгоритм кўплаб соҳаларда қўлланилиб, турли соҳаларда юқори натижаларга эришиш имконини яратади [5,11-21].

Фишер мезонига асосланган алгоритмлар замонавий таҳлил жараёнларида муҳим ўрин тутиб, маълумотларнинг самарали таҳлилини таъминлайди. Ушбу алгоритмлар турли типдаги маълумотлар фазосида қўлланиб, таҳлил жараёнини яхшилашга ёрдам беради. Бу жараёнлар замонавий илмий тадқиқотлар ва амалиётларда катта аҳамиятга эга [6-14].

Қийматли белгилар учун Фишер типидаги мезон асосида ℓ информатив белгилар мажмуасини танлаш алгоритмлари

Фараз қилайлик N - ўлчовли номинал ва қийматли белгилар фазосида қон босимига таъсир этувчи ҳар хил типда берилган дори воситалари ўқув танланмаси $x_{p1}, x_{p2}, \dots, x_{pm_p} \in X_p, p = \overline{1,r}$ берилган бўлсин. Ўқув танланмадаги ҳар бир объект $x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N)$, $i = \overline{1,m_p}$, дори воситаси N - ўлчовли номинал ва қийматли белгилар фазосида берилган. Бу ерда x_{pi}, p-синфнинг i-объекти деб ўқилади, N- объектларни ташкил этувчи белгилар сонини билдирса, m_p- p-синфдаги объектлар сонини билдиради. Демак, $x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N)$ дори воситалари p-синфдаги $i\left(i = \overline{1,m_p}\right)$- объектларни ташкил этади.

Шунингдек, N - ўлчовли белгилар фазосида $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$, бул векторлари киритилади, унинг компонеталари нол ёки бир қиймат қабул қилади. Вектор $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$ нинг компонеталари ёрдамида қайси белги ҳисоблаш ишларида иштирок этаётганлигини ёки этмаётганлиги билинади. Агар $\lambda^j = 1$ бўлса, у ҳолда j-компонент ҳисоблаш ишларида иштирок этади, акс ҳолда $\lambda^j = 0$ бўлса, у ҳолда j-компонент ҳисоблаш ишларида иштирок этмайди.

Таъриф. N-ўлчовли белгилар фазосида $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$ вектор ℓ информатив дейилади, агарда унинг қийматлари $\Lambda^\ell = \left\{ \lambda: \sum_{j=1}^N \lambda^j = \ell\,,\ \lambda^j \in \{0,1\}, j = \overline{1,N} \right\}$ тўпламидан олинса.

Масала 1. Берилган N - ўлчовли қийматли белгилар фазосидан ℓ ўлчовли қийматли белгилар фазосига ўтиш талаб этилсин.

Бу масалани ечишда Фишер типидпги $I(\lambda) = \frac{(a,\lambda)}{(b,\lambda)}$ мезондан фойдаланамиз. Бунинг учун $a = (a^1, a^2, \dots, a^N)$ ва $b = (b^1, b^1, \dots, b^N)$ векторларнинг компоненталарини аниқлаш талаб этилади. Агар масала қийматли белгилар фазосида қаралаётган бўлса, у ҳолда муаммони ҳал қилиш қуйидаги оптимизация масаласини ечишга қаратилган бўлади. Яъни, информатив белгилар мажмуасини танлашнинг оптимизация масаласи қуйидаги кўринишда ифодаланади:

$$\begin{cases} I(\lambda) = \frac{(a,\lambda)}{(b,\lambda)} \to max \\ \Lambda^\ell = \left\{ \lambda: \sum_{j=1}^N \lambda^j = \ell\,,\ \lambda^j \in \{0,1\}, j = \overline{1,N} \right\} \end{cases} \quad (1)$$

бу ерда $a,b \in R^N$ вектор. Бунда мос равишда $a = (a^1, a^2, \dots, a^N)$ ва $b = (b^1, b^1, \dots, b^N)$.

$$a^j = \frac{1}{r-1} \sum_{q=1}^{r} \sum_{p=1}^{r} |\bar{x}_q{}^j - \bar{x}_p{}^j|\,, q \neq p.$$

$$b^j = \frac{1}{(r-1)m_p} \sum_{p=1}^{r} \sum_{i=1}^{m_p} |\bar{x}_p{}^j - x_{pi}{}^j|.$$

Бу ерда $X = \bigcup_{p=1}^r X_p$, $\bar{x}_p$=$(\bar{x}_p^1, \bar{x}_p^2, \dots, \bar{x}_p^N)$ – ўрта объект. Ўрта объект белгилари қиймати қуйидаги формула орқали ҳисобланади:

$$\bar{x}_p^j = \frac{1}{m_p} \sum_{i=1}^{m_p} x_{pi}^j\,, p = \overline{1,r}, j = \overline{1,N}.$$

Алгоритм 1

Юқоридаги оптимизация масаласини ечиш учун ёрдамчи $\varphi(\lambda) = a(b,\lambda) - b(a,\lambda)$ функционални киритиб оламиз.

© А.Х. Нишанов, Ш.У. Актамов, О.Ф. Олламберганов

Фараз қилайлик, φ(λ) векторнинг энг катта компонентасига мос келувчи номери - j1, иккинчиси эса - j2, ва хокоза бўлсин, яъни

$$\varphi_{j1}(\lambda) \geq \varphi_{j2}(\lambda) \geq \varphi_{j3}(\lambda) \geq \ldots \geq \varphi_{jN}(\lambda).$$

μ(λ) деб шундай (μ_1, μ_2, μ_3, ... , μ_N) векторни оламизки, бу ерда

$$\mu_{j1} = \mu_{j2} = \mu_{j3} = \cdots = \mu_{j\ell} = 1$$
$$\mu_{j\ell+1} = \mu_{j\ell+2} = \mu_{j\ell+3} = \cdots = \mu_{jN} = 0$$

Таклиф этилаётган усулда λ векторни танлаш алгоритми қуйидагича:

1-қадам. $\lambda = \{\underbrace{1,1,\ldots,1}_{\ell},0,0,\ldots,0\}$ деб олинади.

2-қадам. $\lambda^* = \mu(\lambda)$ деб ҳисобланади.

3-қадам. Агар $I(\lambda) < I(\lambda^*)$ бўлса, у ҳолда $\lambda = \lambda^*$ деб олинади ва 2-қадамга қайтилади;

4-қадам. λ - мукаммал ечим сифатида қаралади.

Бу алгоритм оптимал ечимга эришиш учун ишлатилади ва ҳар бир қадамда λ вектори янгиланиб боради.

Натижа 1

Информатив белгилар мажмуасини танлаш масаласи тадқиқ этилган бўлиб, бунда 1116 та қон босимига таъсир этувчи ҳар хил турдаги дори-воситалари асосий белгилари жадвал кўринишда олинган. Бунда барча белгилар миқдорий белги бўлган ҳолат тадқиқ этилган. Улар қуйидаги жадвалда қисман берилган:

Жадвал-1

Қон босимига (Гипертензияга қарши) таъсир этувчи дори воситалари ҳақида маълумот										**Синфи**
№	**Дори воситасиноми**	**Фаол модда улуши мг**	**Эрув-чанлик %**	**Қондаги концен-трацияси (Cmax)**	**Тарқалиш ҳажми**	**Тайёрланган таркибларнинг технологик кўрсаткичлари**				
						Сочилувчанлик, 10^{-3} кг/с	**Сочилувчанзичлик, кг/м^3**	**Жигардаги метоболизми %**	**Парчаланиши**	
1	x_1	0,10	96	64	3,50	4,51	20,72	90	7,41	1
2	x_2	0,10	85	78	4,20	7,85	14,36	75	8,56	1
3	x_3	0,50	88	20	4	3,87	25,48	95	5,69	7
4	x_4	0,03	89	72	0,20	16,87	15,64	75	7,96	10
5	x_5	0,20	91	60	3,20	19,20	25,63	85	5,98	10
6	x_6	0,50	93	57	3,90	4,89	48,65	89	6,85	10
.....										
1113	x_{1113}	0,5	76	86	4,01	7,41	13,25	78	9,65	10
1114	x_{1114}	0,25	74	93	4,5	3,41	10,26	60	12,48	10
1115	x_{1115}	0,5	72	94	3,6	4,85	12,48	85	9,85	10

© А.Х. Нишанов, Ш.У. Актамов, О.Ф. Олламберганов

Жадвалдаги асосий кўрсаткичлар қуйидагича:

1. **Дори воситаси номи** – ҳар бир дори учун алоҳида белгиланган идентификатор (x).
2. **Фаол модда улуши** – мг/мл бирлигида кўрсатилган модданинг таркиби.
3. **Эрувчанлик** – фоизда (эритиш қийин ёки осон эканлигини кўрсатади).
4. **Қондаги концентрация (Cmax)** – қон ичидаги максимал концентрацияси.
5. **Тарқалиш зичлиги** – қонда тарқалиш қиймати.
6. **Сочилувчанлик** – дори моддасининг тарқалиш тезлиги (10^{-3} кг/с).
7. **Сочилувчан зичлик** – кг/м^3 бирлигида тарқалишдаги зичлик.
8. **Жигардаги метаболизми** – жигарда парчаланиш фоизи.
9. **Парчаланиши** – молекулаларнинг парчаланиш даражаси ёки тезлиги.

Таклиф этилаётган усулда λ векторни танлаш алгоритми орқали миқдорий белгилардан фойдаланиб, қон босимига таъсир этувчи дори воситаларининг таҳлили учун информатив белгиларнинг оптимал мажмуаси шакллантирилди. Бу белгилар гипертензияга қарши дори воситаларининг самарадорлиги ва уларнинг асосий фармакологик хусусиятларини акс эттиради.(Жадвал-2)

Жадвал-2

Қон босимига (Гипертензияга қарши) таъсир этувчи дори воситалари ҳақида маълумот								
№	Дори воситаси номи	Фаол модда улуши мг	Эрув-чанлик %	Қондаги концен-трацияси (Cmax)	Тарқалиш зичлиги	Тайёрланган таркибларнинг технологик кўрсаткичлари		Синфи
						Сочилувчан зичлик, кг/м^3	Парчалани ши	
1	1	0,10	96	64	3,50	20,72	7,41	1
2	1	0,10	85	78	4,20	14,36	8,56	1
3	7	0,50	88	20	4	25,48	5,69	7
4	10	0,03	89	72	0,20	15,64	7,96	10
5	10	0,20	91	60	3,20	25,63	5,98	10
6	10	0,50	93	57	3,90	48,65	6,85	10
.....								
1113	10	0,5	76	86	4,01	13,25	9,65	10
1114	10	0,25	74	93	4,5	10,26	12,48	10
1115	10	0,5	72	94	3,6	12,48	9,85	10

Синфлаштириш жараёни учун информатив белгиларни танлаш аниқликнинг ортишига олиб келади. Бу каби белгилар асосида дори воситаларини синфлаштириш моделлари юқори аниқлик билан ишлайди, чунки улар дориларнинг фармакологик хусусиятларини яхшироқ акс эттиради ва синфлаштиришни самарали амалга оширишга ёрдам берди. Танланган информатив белгилар мажмуасида бир қанча машҳур алгоритмлар асосида тажриба-синов ўтқазилди. Натижада синфлаштириш жараёни аниқлиги 5-10% га ортиши кузатилди.

Агар объектлар турли типдаги маълумотлар фазосида берилган бўлса у ҳолда Фишер типидаги мезон элементларини топиш талаб этилади. Бунинг учун қуйидаги масала ва тушунча ҳамда белгилашлардан фойдаланамиз.

Масала 2. Берилган N - ўлчовли номинал ва қийматли белгилар фазосидан ℓ ўлчовли белгилар фазосига ўтиш талаб этилсин.

Номинал ва қийматли белгилар фазосида Фишер типидаги мезондан фойдаланиб ℓ информатив белгилар мажмуасини танлаш алгоритми

Номинал белгилар фазосида объектларнинг ўхшашлигини кўрсатувчи катталик $\rho^j(x_{pi,}x_{pq})$ орқали белгиланиб, (2) формула асосида ҳисоблансин, яъни

$$\rho_{pi}^{j}(x_{pi,}x_{pq}) = \begin{cases} 1, & \text{агар} \quad (x_{pi}^{j} - x_{pq}^{j}) = 0; \\ 0, & \text{акс ҳолда} \end{cases} \quad (2)$$

Сонли белгилар фазосида объектларнинг ўхшашлигини кўрсатувчи катталик $\rho^j(x_{pi,}x_{pq})$ орқали белгиланиб, (3) формула асосида ҳисоблансин, яъни

$$\rho_{pi}^{j}(x_{pi,}x_{pq}) = \begin{cases} 1, \text{агар } |x_{pi}^{j} - x_{pq}^{j}| \le \varepsilon^{j} \\ 0, \quad \text{акс ҳолда} \end{cases} \quad (3)$$

$$p = \overline{1,r};\ i \neq q = \overline{1,m_p}; j = \overline{1,N};$$

Ифодаланган (2) ва (3) катталиклар векторнинг параметрлари бўлиб, у қуйидаги $\rho_{pi}(x_{pi,}x_{pq}) = (\rho_{pi}^{1}(x_{pi,}x_{pq}), \rho_{pi}^{2}(x_{pi,}x_{pq}), \dots, \rho_{pi}^{N}(x_{pi,}x_{pq})$ кўринишда ифодаланади. Демак, қаралаётган j-белги номинал бўлса $\rho_{pi}(x_{pi,}x_{pq})$ векторнинг j-белгиси (1) формуладан фойдаланиб ҳисобланади, акс ҳолда, яъни белги сонли бўлса, у ҳолда векторнинг шу белгиси (3) - формула асосида ҳисобланади. Бунда j-белгига мос ε^j – умумий бўсағавий қийматлар синф объектларининг ҳар бир белгиси учун қуйидаги формула орқали амалга оширилади:

$$\varepsilon^{j} = \frac{1}{M-1}\sum_{i=1}^{M-1}|x_i^j - x_{i+1}^j|,$$

бу ерда $j = \overline{1,N}; i = \overline{1,M-1}$; *га тенг.*

Худди шунингдек, синфлар кесимида ε^j – бўсағавий қийматлар

$$\varepsilon_p^j = \frac{1}{m_p - 1}\sum_{i=1}^{m_p-1}|x_{pi}^j - x_{pi+1}^j|,$$

бу ерда $j = \overline{1,N}; p = \overline{1,r}; i = \overline{1,m_p - 1};$ га тенг.

Демак, p синфдаги ихтиёрий иккита x_{pi} *ва* x_{pq} объектлар учун бул векторлар фазосида $\rho_{pi}(x_{pi,}x_{pq}) = (\rho_{pi}^{1}(x_{pi,}x_{pq}), \rho_{pi}^{2}(x_{pi,}x_{pq}), \dots, \rho_{pi}^{N}(x_{pi,}x_{pq}))$ буль векторига эга бўлинади. Бу векторнинг компоненталари қаралаётган белги кесимида иккита объектнинг ўхшашлигини ёки ўхшамаслигини билдиради. Агар $\rho_{pi}^{j}(x_{pi,}x_{pq}) = 1$ бўлса, у ҳолда x_{pi} *ва* x_{pq} объектлар j белги бўйича ўхшаш, акс ҳолда яъни $\rho_{pi}^{j}(x_{pi,}x_{pq}) = 0$ бўлса, у ҳолда шу j параметр бўйича ўхшаш эмастлигини билдиради.

Иккита $\rho_{pi}(x_{pi,}x_{pq})$ ва λ векторларнинг скаляр кўпайтмаси қуйидагича ҳисобланади

$$(\rho_{pi}(x_{pi,}x_{pq}) \cdot \lambda) = \sum_{j=1}^{N}\rho_{pi}^{j}(x_{pi,}x_{pq}) \cdot \lambda^{j}$$

Бу катталик x_{pi} *ва* x_{pq} объектларнинг λ векторга нисбатан ўхшашлик даражасини билдиради. Одатда, бу йиғинди p-синфдаги i-объек, x_{pi} ни, шу p-синфдаги q - объект томонидан баҳолниши деб тушинилади.

Агар $\lambda \in \Lambda^{\ell} = \{\lambda: \sum_{j=1}^{N}\lambda^j = \ell,\ \lambda^j \in \{0,1\}, j = \overline{1,N}\}$ бўлиб ҳамда $(\rho_{pi}(x_{pi,}x_{pq}) \cdot \lambda) = \ell$ бўлса, у ҳолда x_{pi} *ва* x_{pq} объектлар ℓ информатив белгилар фазосида бир-бирига 100% ўхшаш бўлади.

Энди ℓ информатив белгилар фазосида қуйидаги x_{pi} объектни X_p синфнинг барча объектлари томонидан ўртача баҳосини аниқлаш талаб этилсин. Бу баҳони берувчи мезонни $\boldsymbol{\Gamma}(\boldsymbol{\lambda}, x_{pi}, X_p)$ орқали белгилайлик. Бу мезонни маъноси ℓ информатив белгилар фазосида x_{pi} объектни X_p синфнинг барча

объектлари томонидан ўртача баҳоланиши тушунилади ва унинг математик ифодаланиши қуйидагича бўлади. Айрим адабиётларда бу катталикни x_{pi} объектнинг ахборот ўлчови деб тушинилади. Чунки бу катталик x_{pi} объектнинг X_p синфдаги муҳимлик даражасини билдиради.

$$\boldsymbol{\Gamma}(\boldsymbol{\lambda}, x_{pi}, X_p) = \frac{1}{m_p - 1}\sum_{q=1}^{m_p}(\rho_{pi}(x_{pi,}x_{pq}), \boldsymbol{\lambda}), i \neq q,$$

Бу ифоданинг маноси X_pсинфнинг барча объектлари томонидан $x_{pi,}$объектнинг ахборот ўлчовини аниқлаб беради.

Шунингдек, X_p синфнинг барча объектларининг ахборот ўлчовини аниқлаш мезони $\boldsymbol{\Gamma}(\boldsymbol{\lambda}, X_p)$ кўринишда ўз ифодасини топади ва қуйидагича ҳисобланади:

$$\boldsymbol{\Gamma}(\boldsymbol{\lambda}, X_p) = \frac{1}{m_p}\sum_{i=1}^{m_p}\boldsymbol{\Gamma}(\boldsymbol{\lambda}, x_{pi}, X_p) = \frac{1}{m_p(m_p-1)}\sum_{i=1}^{m_p}\sum_{q=1}^{m_p}(\rho_{pi}(x_{pi,}x_{pq}), \boldsymbol{\lambda}), i \neq q. \quad (4)$$

Демак, турли типдаги маълумотлар фазосида тимсолларни (4) ахборот ўлчовини ҳисоблаш имкониятига эга бўлинди.

Шунингдек, X_p синфнинг барча объектларининг ахборот ўлчовини аниқлаш мезони $\boldsymbol{I}(\boldsymbol{\lambda}, X_p)$ кўринишда ўз ифодасини топади ва қуйидагича ҳисобланади:

$$\Gamma(\lambda, X_p) = \frac{1}{m_p}\sum_{i=1}^{m_p} I(\lambda, x_{pi}, X_p) = \frac{1}{m_p(m_p-1)}\sum_{i=1}^{m_p}\sum_{q=1}^{m_p}(\rho_{pi}(x_{pi,}x_{pq}), \lambda), i \neq q. \quad (5)$$

Албатта, $\lambda \in \Lambda^{\ell} = \{\lambda: \sum_{j=1}^{N}\lambda^j = \ell,\ \lambda^j \in \{0,1\}, j = \overline{1,N}\}$ фазода $\boldsymbol{\Gamma}(\boldsymbol{\lambda}, X_p)$ баҳоси қанча катта бўлса, у ҳолда X_p синфнинг барча объектларининг ўзаро ўхшашлик даражаси шунчалик катта бўлади. Демак, бу объектлар ℓ информатив белгилар фазосида бир-бирига жуда яқин, ўхшаш деб ҳисобланади. Бир синф объектлари учун (5) нинг қиймати қанча катта бўлса, шунча ўша синф объектлари ўзаро бир – бирига ўхшаш деб тушинилади. Демак, (5) эвристик мезон бўлиб, унинг қийматлари синфни ҳолатини баҳолаб беради.

Яна битта мезон киритайлик, унда турли типдаги маълумотлар фазосида X_p синф объектларини X_q $(p \neq q)$ объектлари томонидан ўхшашлик даражалари баҳолансин ва уни $\boldsymbol{\Gamma}(\boldsymbol{\lambda}, X_p, X_q), (p \neq q)$ орқали белгилаб қуйидагича ҳисоблансин:

$$\boldsymbol{\Gamma}(\boldsymbol{\lambda}, X_p, X_q) = \frac{2}{m_p m_q}\sum_{i=1}^{m_p}\sum_{k=1}^{m_q}(\rho_{pi}(x_{pi,}x_{qk}), \boldsymbol{\lambda}), p \neq q. \quad (6)$$

Одатда, (6) нинг қийматининг кичиклиги иккита синф объектларининг ўзаро ўхшамаслик даражасини беради.. Демак, функционалнинг қиймати қанчалик кичик бўлса бу синфлар бир биридан шунчалик узоқлашадеб тушинилади.

Қуйидаги оптимизация масаласини қарайлик

$$\begin{cases} I(\lambda) = \frac{\Gamma(\lambda, X_p)}{\Gamma(\lambda, X_p, X_q)} \to max \\ \Lambda^{\ell} = \{\lambda: \sum_{j=1}^{N}\lambda^j = \ell,\ \lambda^j \in \{0,1\}, j = \overline{1,N}\} \end{cases} \quad (7)$$

Демак, мақсад функциясининг суръати қиймати синф объектларининг ўзаро ўхшашлик даражасини билдиради, шунинг учун ҳам у қанча катта ва махражи синфлар аро объектларнин ўхшашлик даражаларини билдиради , у қанчалик кичик бўлса, танланган ℓ информатив белгилар мажмуаси шунчалик мукаммал топилган деб тушинилади. Албатта, (7) эвристик мезон ва (7) масаланинг ечимини топиш билан шуғилланамиз.

Қаралаётган (5) ва (6) функционални ҳисоблаётганда бу функционаллар λ векторга боғлиқ бўлмаганлиги учун уларни $a = (a^1, a^2, \dots, a^N)$ ва $b = (b^1, b^1, \dots, b^N)$ векторлар компоненталари орқали қуйидагича ифодалаш мумкин.

$$a^j = \frac{1}{m_p(m_p-1)} \sum_{i=1}^{m_p} \sum_{k=1}^{m_p} \rho_{pi}^j(x_{pi,} x_{pk}), k \neq i, j = \overline{1,N}.$$

Бу формула p-синф ёки X_p синфдаги объектларнинг ички ўзаро алоқасини ёки ўхшашлик даражаларини ўлчайди. Бу ерда m_p, X_p синфга тегишли объектлар сони. ρ_{pi}^j : j-чи атрибут учун икки элемент орасидаги ўзаро алоқа ёки ўлчов функцияси. Бу функция объектлар орасидаги яқинлик ёки фарқни ўлчаш учун фойдаланилади. x_{pi} *ва* x_{pk}: p-синфдаги i - ва k- объектлар. $k \neq i$ шарти – бу ҳар бир элемент ўзидан бошқа барча элементлар билан таққосланишини билдиради.

Бу формула синф ичидаги объектларнинг бир-бирига яқинлигини ўртача даражада ўлчаш учун ишлатилади. Масалан, агар p-синфда бир гуруҳ одамлар бўлса, уларнинг ўзаро ўхшашликлари ёки фарқларини ҳисоблаш учун фойдаланилади.

Шунингдек,

$$b^j = \frac{2}{m_p m_q} \sum_{i=1}^{m_p} \sum_{k=1}^{m_q} \rho_{pi}^j(x_{pi,} x_{qk}), \quad q \neq p, j = \overline{1,N}.$$

Бу формула икки синф p- ва q-синфлар орасидаги ўзаро алоқани ёки ўхшашалик даражаларини баҳолаш учун мўлжалланган. Бу ерда m_p ва m_q мос равишда X_p ва X_q синфлардаги объектлар сони, ҳамда ρ_{pi}^j j-чи атрибут бўйича элементлар орасидаги алоқа ёки масофа ўлчови, x_{pi} ва x_{qk} мос равишда p- ва q-синфлардаги объектлар.

Бу формула иккита турли синф объекталри ўртасидаги ўртача ўзаро ўхшашлик даражаларини ёки алоқани аниқлайди. Мақсад – турли синф объектлари ўртасидаги фарқларни ёки ўхшашликларни баҳолайди.

Юқоридаги келтирилган формулалар (7), (8), (9) лардан фойдаланиб (1) формулани янги қийматли ва номиқдорий белгилар кесимида қуйидагича ифодаланади

$$\begin{cases} I(\lambda) = \frac{(a,\lambda)}{(b,\lambda)} \to max \\ \Lambda^{\ell} = \left\{ \lambda : \sum_{j=1}^{N} \lambda^j = \ell, \ \lambda^j \in \{0,1\}, j = \overline{1,N} \right\} \end{cases} (1')$$

Бу ерда (a,λ) ва (b,λ) - вазнли йиғиндилар ёки чизиқли комбинациялар, λ-буль вектори ва $\Lambda^{\ell} = \left\{ \lambda : \sum_{j=1}^{N} \lambda^j = \ell, \ \lambda^j \in \{0,1\}, j = \overline{1,N} \right\}$ максимизация шарти.

Бу функциянинг мақсади – синфлар ичидаги ва синфлар ўртасидаги алоқалар нисбатини максимал даражада фарқ қилишдир. Масалан, кластерлаш ёки тоифалаш масаласида, элементларни шундай гуруҳлаш керакки, гуруҳлар ичидаги элементлар бир-бирига яқин, аммо турли гуруҳлар элементлари бир-биридан узоқ бўлсин.

Бу формулалар маълумотларни кластерлашда, яъни элементларни гуруҳларга ажратишда қўлланилиши мумкин. Улар тоифа ичидаги ва тоифа ўртасидаги ўзаро алоқаларни ўлчашга ёрдам беради. Бундай формулалар кластерлар орасидаги аниқ фарқларни ёки ўхшашликларни топиш учун жуда фойдали.

Бу оптимизация масаласини ечими иккита синф ўртасидаги ўхшашликни камайтирадиган ва бир пайтни ўзида синф объектлари ўртасидаги ўхшашликни оширадиган информатив белгилар мажмуасини танлаш имконини беради. Қуйида (1′) оптимизация масаласини ечимини топишнинг алгоритми келтирилади.

Алгоритм 2

Юқоридаги оптимизация масаласини ечиш учун ёрдамчи $\psi(\lambda) = a(b,\lambda) - b(a,\lambda)$ функционални киритиб олинади.

Фараз қилайлик, $\psi(\lambda)$ векторнинг энг катта компонентасига мос келувчи номери - j1, иккинчиси эса - j2, ва хокоза бўлсин, яъни

$$\psi_{j1}(\lambda) \geq \psi_{j2}(\lambda) \geq \psi_{j3}(\lambda) \geq \ldots \geq \psi_{jN}(\lambda).$$

© А.Х. Нишанов, Ш.У. Актамов, О.Ф. Олламберганов

Таклиф этилаётган усулда λ векторни танлаш алгоритми қуйидагича:

1-қадам. $\lambda = \{\underbrace{1,1,\dots,1}_{\ell},0,0,\dots,0\}$ деб олинади.

2-қадам. $a = (a^1, a^2, \dots, a^N)$ ва $b = (b^1, b^1, \dots, b^N)$ векторлар компоненталарини аниқлаш талаб этилади. Бунда a^j ва b^j лар қуйидаги формулалар асосида ҳисобланади:

$$a^j = \frac{1}{m_p(m_p-1)}\sum_{i=1}^{m_p}\sum_{k=1}^{m_p}\rho_{pi}^j(x_{pi,}x_{pk}), k \neq i,j = \overline{1,N},$$

$$b^j = \frac{2}{m_p m_q}\sum_{i=1}^{m_p}\sum_{k=1}^{m_q}\rho_{pi}^j(x_{pi,}x_{qk}), \quad q \neq p, j = \overline{1,N}.$$

3-қадам. Ҳар бир λ учун (a,λ) ва (b,λ) скаляр кўпайтма ҳисобланади.

4-қадам. $I(\lambda) = \frac{(a,\lambda)}{(b,\lambda)}$ функционали қийматини топилади.

5-қадам. Агар янги ҳисобланган $I(\lambda)$ қиймати олдинги ҳисобланган максимум қийматдан катта бўлса, унда максимал қийматни ва λ векторини янгиланади. $\psi(\lambda) = a(b,\lambda) - b(a,\lambda)$ функционали қийматини қайта ҳисоблаш амалга оширилади. 3- ва 4-қадамлар такрорланади, то ҳолат максимал қийматга етгунча.

6-қадам. λ - мукаммал ечим сифатида қаралади.

Алгоритм оптимал ечимга эришиш учун ишлатилади ва ҳар бир қадамда λ вектори янгиланиб боради.

Натижа 2

Информатив белгилар мажмуасини танлаш масаласи тадқиқ этилган бўлиб, бунда 1116 та қон босимига таъсир этувчи ҳар хил турдаги дори-воситалари асосий белгилари жадвал кўринишда олинган. Бунда 9 та белги миқдорий белги ва битта номинал белги бўлган ҳолат тадқиқ этилган. Улар қуйидаги жадвалда қисман берилган:

Жадвал-3

Қон босимига (Гипертензияга қарши) таъсир этувчи дори воситалари ҳақида маълумот												
№	Дори воситасиноми	Фаол модда улуши %	Эрув-чанлик %	Ранги	Ишлаб чиқариш шакли	Қондаги концен-трацияси (Cmax)	Таркалиш зичлиги	Тайёрланган таркибларнинг технологик кўрсаткичлари				Синфи
								Сочилувчанлик, 10^{-3} кг/с	Сочилувчан зичлик, кг/м3	Жигардаги мотеболизми %	Парчаланиши	
1	x_1	0,10	96	сариқ	10	64	3,50	4,51	20,72	90	7,41	1
2	x_2	0,10	85	оқ	20	78	4,20	7,85	14,36	75	8,56	1
3	x_3	0,50	88	оч сариқ	25	20	4	3,87	25,48	95	5,69	7
4	x_4	0,03	89	оқ	50	72	0,20	16,87	15,64	75	7,96	10
5	x_5	0,20	91	оқ	15	60	3,20	19,20	25,63	85	5,98	10

© А.Х. Нишанов, Ш.У. Актамов, О.Ф. Олламберганов

6	x_6	0,50	93	оқ	25	57	3,90	4,89	48,65	89	6,85	10
.....												
1113	x_{1113}	0,5	76	оқ	25	86	4,01	7,41	13,25	78	9,65	10
1114	x_{1114}	0,25	74	кулранг	14	93	4,5	3,41	10,26	60	12,48	10
1115	x_{1115}	0,5	72	оқ	28	94	3,6	4,85	12,48	85	9,85	10

Таклиф этилаётган усулда λ векторни танлаш алгоритми орқали миқдорий ва номинал белгилардан фойдаланиб, қон босимига таъсир этувчи дори воситаларининг таҳлили учун информатив белгиларнинг оптимал мажмуаси шакллантирилди. Бунда номинал белги информатив белгилар мажмуасини танлашда қатнашиши белгилаб қўйилди. Алгоритм натижаси қуйидаги жадвалда келтирилган. (Жадвал-4)

Жадвал-4

Қон босимига (Гипертензияга қарши) таъсир этувчи дори воситалари ҳақида маълумот								
№	Дори воситаси номи	Фаол модда улуши мг	Эрув-чанлик %	Ранги	Қондаги концен-трацияси (Cmax)	Тайёрланган таркибларнинг технологик кўрсаткичлари		Синфи
						Сочилувчанзичлик, кг/м3	Парчаланиши	
1	x_1	0,10	96	сариқ	64	20,72	7,41	1
2	x_2	0,10	85	оқ	78	14,36	8,56	1
3	x_3	0,50	88	оч сариқ	20	25,48	5,69	7
4	x_4	0,03	89	оқ	72	15,64	7,96	10
5	x_5	0,20	91	оқ	60	25,63	5,98	10
6	x_6	0,50	93	оқ	57	48,65	6,85	10
.....								
1113	x_{1113}	0,5	76	оқ	86	13,25	9,65	10
1114	x_{1114}	0,25	74	кулранг	93	10,26	12,48	10
1115	x_{1115}	0,5	72	оқ	94	12,48	9,85	10

© А.Х. Нишанов, Ш.У. Актамов, О.Ф. Олламберганов

Жадвалдаги асосий кўрсаткичлар қуйидагича:

Дори воситаси номи – ҳар бир дори учун алоҳида белгиланган идентификатор (x).

Фаол модда улуши – мг/мл бирлигида кўрсатилган модданинг таркиби.

Эрувчанлик – фоизда (эритиш қийин ёки осон эканлигини кўрсатади).

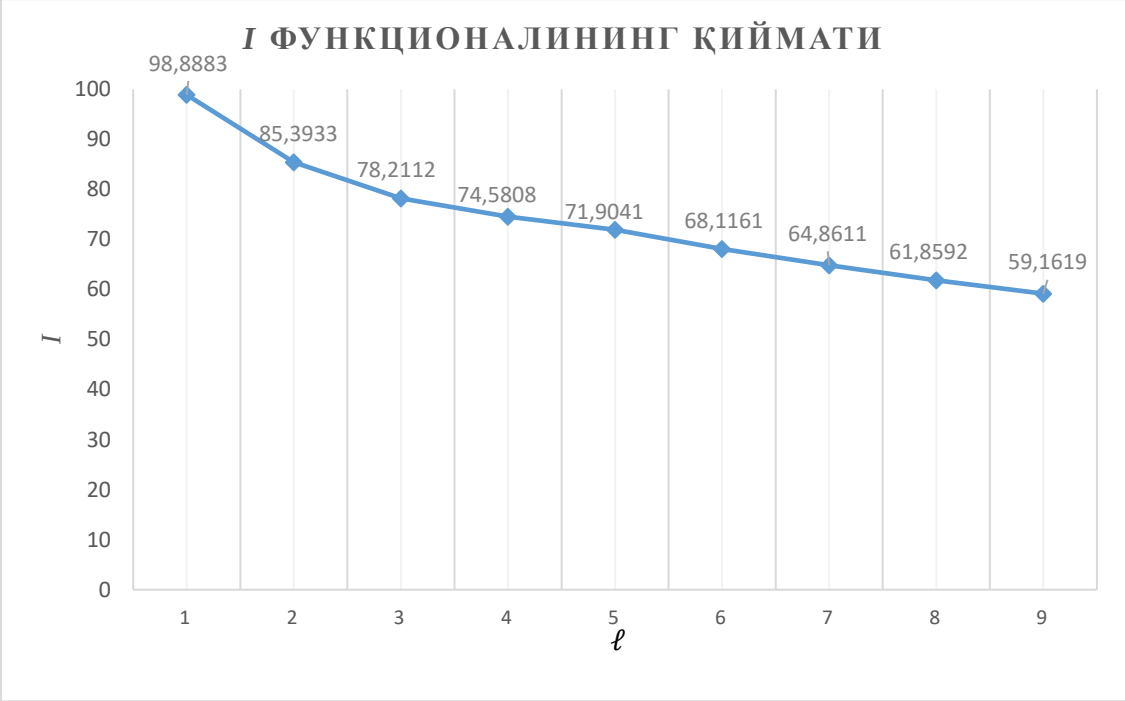
Ранги – ҳар бир дорининг рангини ифодалайди.

Қондаги концентрация (Cmax) – қон ичидаги максимал концентрацияси.

Сочилувчан зичлик – кг/м3 бирлигида тарқалишдаги зичлик.

Парчаланиши – молекулаларнинг парчаланиш даражаси ёки тезлиги.

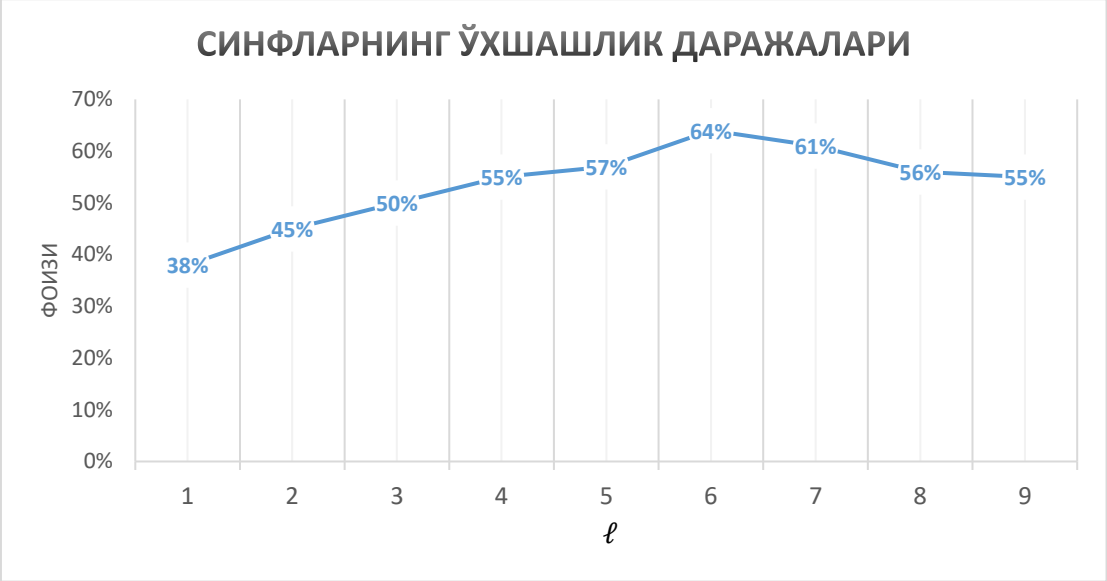
Фишер мезонига асосланган информатив белгиларни танлаш алгоритмида асосида ҳисобланган I фунционали қийматининг ℓ га боғлиқ ўзгариш графиги келтирилган(1-расм).



1-расм. ℓ нинг қийматига мос I функционалининг ўзгариши

Информатив белгилар мажмуасини танлашга асосланган алгоритми асосида олинган натижалар асосида ўқув танланмадаги синфлар объектларнинг ўзаро ўхшашлик даражалари графиги келтирилган(2-расм).

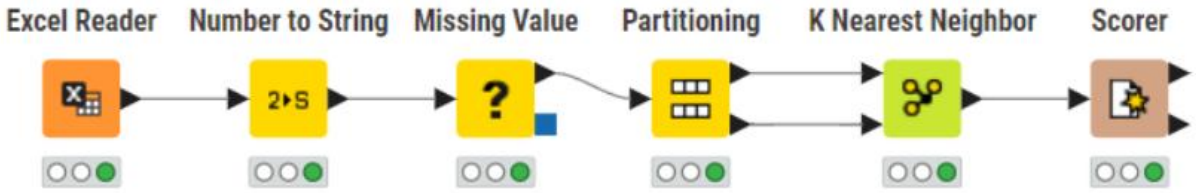
© А.Х. Нишанов, Ш.У. Актамов, О.Ф. Олламберганов



2-расм. ℓ нинг қийматига мос синфларнинг ўхшашлик даражаларининг ўзгариши

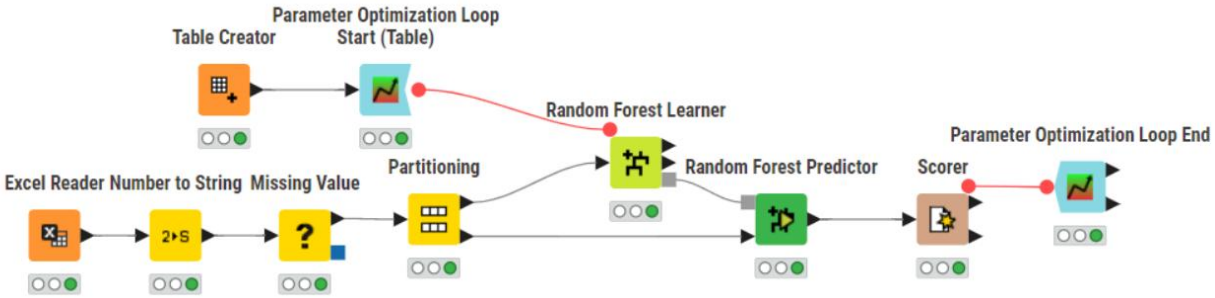
Тажриба-синов

Олинган натижалар **KNIME** муҳитида қурилган моделлар орқали тажриба-синовдан ўтказилди. KNIME муҳитидаги KNN алгоритмининг модели 1-расмда келтирилган.



1-расм. KNN алгоритмининг KNIME платформасидаги модели

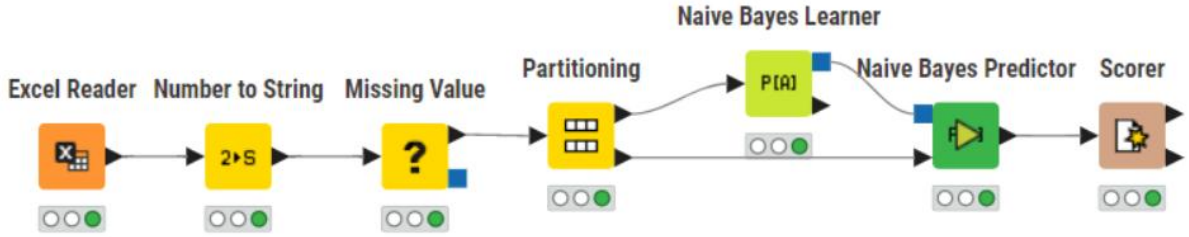
KNIME муҳитидаги Random Forest алгоритмининг модели 2-расмда келтирилган.



2-расм. Random Forest алгоритмининг KNIME платформасидаги модели

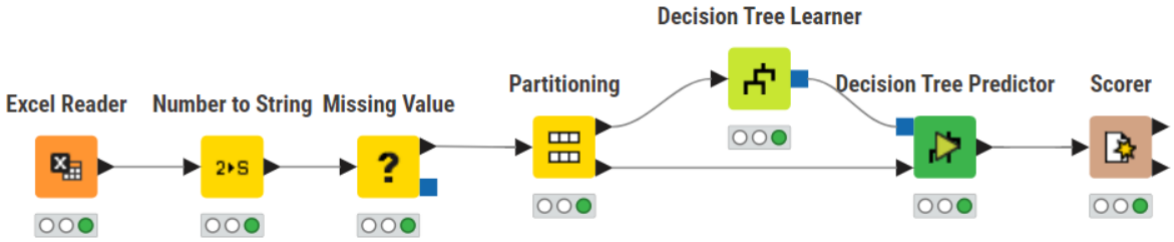
KNIME муҳитидаги Naive Bayes алгоритмининг модели 3-расмда келтирилган.

© А.Х. Нишанов, Ш.У. Актамов, О.Ф. Олламберганов



3-расм. Naive Bayes алгоритмининг KNIME платформасидаги модели

KNIME муҳитидаги Decision Tree алгоритмининг модели 4-расмда келтирилган.



4-расм. Decision Tree алгоритмининг KNIME платформасидаги модели

Таклиф этилган назарий тадқиқотлар, алгоритмлар асосида юқорида баён этилган масала ечилган бўлиб қон босимига (Гипертензияга қарши) таъсир этувчи дори воситаларини тўғри синфлаштириш бўйича ҳар бир синф кесимида ва алгоритмлар кесимида 5-жадвалда ўз аксини топган.

Жадвал-5

Алгоритмлар тахлилидан олинган натижалар

№	K Nearest Neighbor алгоритми	Random Forest алгоритми	Naive Bayes алгоритми	Decision Tree алгоритми
1-синф	95%	94%	100%	97%
2-синф	91%	100%	93%	91%
3-синф	87%	93%	87%	80%
4-синф	96%	95%	100%	89%
4-синф	93%	92%	85%	88%
5-синф	97%	90%	95%	87%
6-синф	96%	94%	96%	89%
7-синф	100%	100%	88%	86%
8-синф	94%	94%	100%	92%
9-синф	100%	85%	80%	100%
10-синф	96%	95%	100%	97%

Ўқув танланма объектларини кластерлаш ва синфлаштириш масалари ечими ҳамда натижалар олдинги ишларда келтирилган [12]. Информатив белгилар мажмуасини танлаш амалга оширилгандан кейин синфлаштириш аниқлиги 5-13% га кўтарилди.

Хулоса

Фишер типидаги мезонга асосланган информатив белгиларни танлаш алгоритми асосан класслар ўртасидаги фарқни аниқлашга ва уларнинг объектларининг самарали тақсимотини таъминлашга қаратилган. Ушбу алгоритм класслар ўртасидаги фарқларни юқори даражада сақлаб қоладиган белгиларни танлайди. Ушбу усул классификациядаги аниқликни оширишга ёрдам беради.

Фишер типидаги мезонга асосланган информатив белгиларни танлаш алгоритми ҳақида қуйидаги хулосаларни келтириш мумкин:

Мазкур алгоритм турли типдаги маълумотлар фазосидаги қон босимига таъсир этувчи дори воситалари учун информатив белгиларни танлашга йўналтирилган.

Алгоритм Фишер мезони асосида белгиланган мақсад функциясини оптималлаштиришга қаратилган бўлиб, бу мезон объектларнинг ўхшашлик даражасини баҳолайди.

Векторларнинг компоненталарини ҳисоблаш ва оптимал ечимни топиш алгоритмнинг асосий қадамларини ташкил этади.

Мақоладаги (5) ва (6) функционаллар синф объектларининг ўхшашлик даражасини аниқлаш учун фойдаланиладиган эвристик мезонлардир.

Оптимизация масаласи информатив белгилар мажмуасини танлашда синф объектларининг ўзаро ўхшашлик ва фарқларини ҳисобга олишни таъминлайди.

Фойдаланилган адабиётлар

1. Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics, 7(2), 179-188.
2. Fukunaga, K. (1990). Introduction to Statistical Pattern Recognition (2nd ed.). Academic Press.
3. Duda, R. O., Hart, P. E., & Stork, D. G. (2001). Pattern Classification (2nd ed.). Wiley-Interscience.
4. Cover, T. M., & Thomas, J. A. (2006). Elements of Information Theory (2nd ed.). Wiley-Interscience.
5. Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(1), 4-37.
6. Theodoridis, S., & Koutroumbas, K. (2009). Pattern Recognition (4th ed.). Academic Press.
7. Nishanov A. Ruzibaev O. Tran N. Modification of decision rules 'ball Apolonia' the problem of classification // 2016 International Conference on Information Science and Communications Technologies, ICISCT 2016 (2016),doi: 10.1109/ICISCT.2016.7777382.
8. Nishanov A., Saidrasulov Sh., Babadjanov E. Analysis of Methodology Of Rating Evaluation Of Digital Economy And E-Government Development In Uzbekistan // International Journal Of Early Childhood Special Education, Volume:14 Issue:2 2022, p. 2447-2452, doi: 10.9756/INT-JECSE/V14I2.230.
9. Nishanov A., Ruzibaev O., Chedjou J. C., Kyamakya K., Abhiram, Kolli, De Silva, Djurayev G., Khasanova M. Algorithm for the selection of informative symptoms in the classification of medical data // Developments Of Artificial Intelligence Technologies In Computation And Robotics, Volume:12, 2020, p. 647-658, doi:10.1142/9789811223334_0078
10. Nishanov AKh, Turakulov AKh, Turakhanov KhV. Reshaiushchee pravilo dlia klassifikatsii patologii zritel'noĭ sistemy [A decisive rule in classifying diseases of the visual system]. Med Tekh. 1999 Jul-Aug;(4):16-8. Russian. PMID: 10464756.
11. Nishanov, A., Akbarova, M., Tursunov, A., Ollamberganov, F., & Rashidova, D. (2024). Clustering algorithm based on object

© А.Х. Нишанов, Ш.У. Актамов, О.Ф. Олламберганов

similarity. Journal of Mathematics, Mechanics and Computer Science, 123(3), 108–120. https://doi.org/10.26577/JMMCS2024-v123-i3-4

12. Ollamberganov, F. (2024). Объектларнинг муҳимлилик даражаларига мос кластерлаш ва синфлаштириш алгоритмлари. Digital transformation and artificial intelligence, 2(4), 113-125.
13. Nishanov, Akhram & Djuraevb, Gulomjon & Khasanova, Malika & Saparov, Saidqul & Zaripov, Fazilbek. (2023). Algorithm of diagnostics of medical datas based on symptom complexes. 34. 10.1117/12.2669449.
14. Nishanov, A.K., Juraev, G.P., Khasanova, M.A., Zaripov, F.M., Saparov, S.X. (2023). Algorithm for the Classification of Coronary Heart Disease Based on the Use of Symptom Complexes in the Cardiovascular Environment. In: High-Performance Computing Systems and Technologies in Scientific Research, Automation of Control and Production. HPCST 2022. Communications in Computer and Information Science, vol 1733. Springer, Cham. https://doi.org/10.1007/978-3-031-23744-7_12

© *А.Х. Нишанов, Ш.У. Актамов, О.Ф. Олламберганов*