



O'QITISHDA QARORLAR DARAXTINI HOSIL QILISHNI O'QUVCHILARGA YANGI TEXNOLOGIYALARGA ASOSLANGAN HOLDA O'RGATISH

Xajiboyeva Surayyo¹
Allaberganova Gulnoza²
Allaberganova Nasiba³
Babajanov Umar⁴

¹⁻²⁻³⁻⁴Urganch davlat universiteti Axborot texnologiyalari kafedrası o'qituvchisi
<https://doi.org/10.5281/zenodo.6579861>

Annotatsiya. Ushbu maqolada qarorlar daraxtini hosil qilishni o'quvchilarga yangi texnologiyalarga asoslangan holda o'rgatish masalasi ko'rib chiqilgan.

Аннотация. В данной работе изучается , как научить студентов создавать дерево решений на основе новых технологий.

Annotation. This article addresses the issue of teaching students how to create a decision tree based on new technologies.

Kalit so'zlar: Entropiya, Ma'lumot yutug'i, qarorlar daraxti, qaror tugunlari.

Ключевые слова: Энтропия, Прирост информации, Дерево решений, Узлы решений.

Key words: Entropy, Information Gain, Decision Tree, Decision Nodes.

Qaror daraxti tasniflash regressiya modellarini daraxt strukturasi shaklida yaratadi. U ma'lumotlar to'plamini kichik to'plamlarga ajratadi va shu bilan birga tegishli qarorlar daraxti bosqichma-bosqich ishlab chiqiladi. Yakuniy natija - qaror tugunlari va barg tugunlari bo'lgan daraxt . Qaror tugunida (masalan, Holat) ikki yoki undan ortiq atribut mavjud (masalan, quyoshli, bulutli va yomg'irli). Barg tugunlari (masalan, O'ynash) tasnif yoki qarorni ifodalaydi. Daraxtdagi eng yaxshi bashoratchiga mos keladigan eng yuqori qaror tuguniga ildiz tugunlari deyiladi . Qaror daraxtlari ham kategorik, ham raqamli ma'lumotlarni qayta ishlashga qodir.

Quyidagi ma'lumotlar to'plamini ko'rib chiqamiz. Bizga 4 ta atribut- holat, temperatura, namlik va shamol darajalari berilgan. Shu asosda futbol o'ynash mumkinmi yoki mumkin emasmi shuni aniqlashimiz lozim.

<i>Holat</i>	<i>Temperatura</i>	<i>Namlik</i>	<i>Shamol</i>	<i>Futbol o'ynash</i>
Yomg'irli	Issiq	Yuqori	Yolg'on	Yo'q
Yomg'irli	Issiq	Yuqori	Rost	Yo'q
Bulutli	Issiq	Yuqori	Yolg'on	Ha
Quyoshli	Iliq	Yuqori	Yolg'on	Ha
Quyoshli	Sovuq	O'rtacha	Yolg'on	Ha
Quyoshli	Sovuq	O'rtacha	Rost	Yo'q

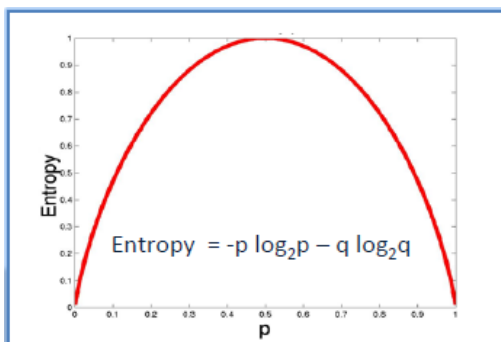


Bulutli	Sovuq	O'rtacha	Rost	Ha
Yomg'irli	Iliq	Yuqori	Yolg'on	Yo'q
Yomg'irli	Sovuq	O'rtacha	Yolg'on	Ha
Quyoshli	Iliq	O'rtacha	Yolg'on	Ha
Yomg'irli	Iliq	O'rtacha	Rost	Ha
Bulutli	Iliq	Yuqori	Rost	Ha
Bulutli	Issiq	O'rtacha	Yolg'on	Ha
Quyoshli	Iliq	Yuqori	Rost	Yo'q

JR Quinlan tomonidan ID3 deb nomlangan qarorlar daraxtlarini qurish uchun asosiy algoritm hisoblanadi. ID3 qarorlar daraxtini yaratish uchun *Entropiya* va *Ma'lumot yutug'idan* foydalanadi.

Entropiya

Qaror daraxti ildiz tugunidan yuqoridan pastga qurilgan va ma'lumotlarni o'xshash qiymatlarga ega (bir hil) misollarni o'z ichiga olgan kichik to'plamlarga bo'lishni o'z ichiga oladi. ID3 algoritmi namunaning bir xilligini hisoblash uchun entropiyadan foydalanadi. Agar namuna butunlay bir hil bo'lsa, entropiya nolga teng, agar namuna teng bo'lingan bo'lsa, u bitta entropiyaga ega (1-rasm).



$$\text{Entropy} = -0.5 \log_2 0.5 - 0.5 \log_2 0.5 = 1$$

1-rasm. Entropiya grafigi.

a) Bitta atributning chastotalar jadvalidan foydalangan holda entropiya:

$$E(S) = \sum_{i=1}^c -p_i \log_2 p_i$$

Futbol o'ynash	
Ha	Yo'q
9	5

$$E(\text{Futbol o'ynash}) = E(5,9) = E(0.36,0.64) = -(0.36 \log_2 0.36) - (0.64 \log_2 0.64) = 0.94$$

1-jadval

b) Ikki atributning chastotalar jadvalidan foydalangan holda entropiya:





		Futbol o'ynash		
		Ha	Yo'q	
Holat	Quyoshli	2	3	5
	Bulutli	4	0	4
	Yomg'irli	3	2	5
				14

2-jadval

$$E(T, X) = \sum_{c \in X} P(c)E(c)$$

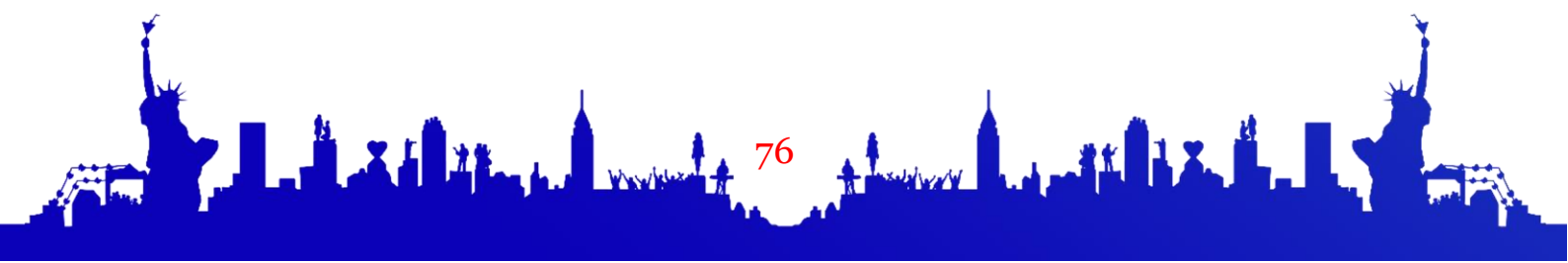
$$E(\text{Futbol o'ynash}, \text{Holat}) = P(\text{Quyoshli}) * E(2,3) + P(\text{Bulutli}) * E(4,0) + P(\text{Yomg'irli}) * E(2,3) = (5/14) * 0.971 + (4/14) * 0.0 + (5/14) * 0.971 = 0.693$$

Ma'lumot yutug'i

Qarorlar daraxtini yaratish - bu eng yuqori ma'lumot (ya'ni, eng bir hil bo'laklar) qaytaradigan atributni topishdir. Ma'lumot yutug'ini hisoblash quyidagi shaklda amalga oshiriladi.

1-qadam : Entropiyani hisoblang.

$$E(\text{Futbol o'ynash}) = E(5,9) = E(0.36,0.64) = -(0.36 \log_2 0.36) - (0.64 \log_2 0.64) = 0.94$$





2-qadam : Keyin ma'lumotlar to'plami turli bo'laklarga bo'linadi. Har bir bo'lak uchun bitta atributli va ikkita atributli entropiya hisoblanadi va ular asosida ma'lumot yutug'i hisoblanadi.

		Futbol o'ynash	
		Ha	Yo'q
Shamol	Yolg'on	6	2
	Rost	3	3
Ma'lumot yutug'i= 0.048			
		Futbol o'ynash	
		Ha	Yo'q
Holat	Quyoshli	2	3
	Bulutli	4	0
	Yomg'irli	3	2
Ma'lumot yutug'i= 0.247			

		Futbol o'ynash	
		Ha	Yo'q
Temperatura	Issiq	2	2
	Iliq	4	2
	Sovuq	3	1
Ma'lumot yutug'i= 0.029			

		Futbol o'ynash	
		Ha	Yo'q
Namlik	Yuqori	3	4
	O'rtacha	6	1
Ma'lumot yutug'i= 0.152			

3-jadval

$$Gain(T, X) = Entropy(T) - Entropy(T, X)$$

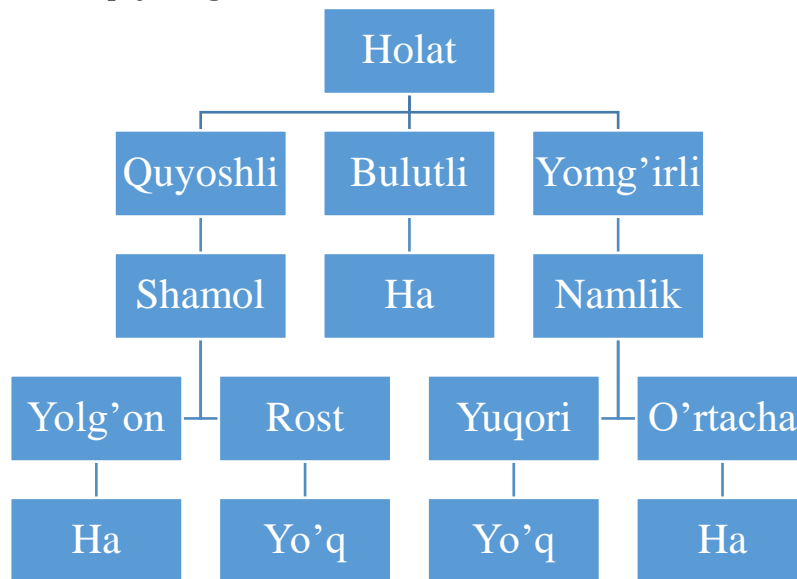
$$G(Futbol o'ynash, Holat) = E(Futbol o'ynash) - E(Futbol o'ynash, Holat) \\ = 0.940 - 0.693 = 0.247$$

3-qadam : Qaror tugunlari sifatida eng katta ma'lumotga ega bo'lgan atributni tanlang, ushbu holatda 3-jadvalning birinchi bo'laki tanlab olinadi. Ma'lumotlar to'plamini bo'laklarga bo'ling va har bir bo'lakda bir xil jarayonni takrorlang.





4-qadam : Yuqorida ma'lumotlar asosida qarorlar daraxtini hosil qilamiz.
Uning ko'rinishi quyidagicha bo'ladi.



Qaror daraxtini ildiz tugunidan barg tugunlariga birma-bir xaritalash orqali osongina qoidalar to'plamiga aylantirish mumkin.

Foydalanilgan adabiyotlar:

1. S.Komolov, Sh.Raxmatov, "Sun'iy intellekt asoslari" 95-105b.
2. Sh.A.Sadullayeva, D.F.Yusupov, F.Yusupov, "Sun'iy intellekt va neyronto'rli texnologiyalar" 9-17b
3. https://www.saedsayad.com/decision_tree.htm

