

**TAKSONOMIYA USULLARI VA TAJRIBA BERILGANLARNI
BOSHLANG'ICH TAHLILI**

Tojimatov Israiljon Nurmamatovich

Farg'ona davlat universiteti amaliy matematika va informatika kafedrasida katta o'qituvchisi

israiltojimatov@gmail.com

Qo'shaqova Dilshoda Robiljon qizi

Farg'ona davlat universiteti 3-kurs talabasi

qoshaqovadilshodahon@gmail.com

Annotatsiya: Taksonomiya ilm-fan sohasida tizimlashtirishning muhim vositasi hisoblanadi. U biologik turlardan tortib, axborot texnologiyalari, iqtisodiyot, pedagogika va boshqa ko'plab sohalarda keng qo'llaniladi. Ushbu maqola taksonomiya usullari va tajriba berilgan ma'lumotlarni boshlang'ich tahlil qilish jarayonini tahlil qiladi. Taksonomiya tushunchasi, uning tarixi va amaliy sohalarda qo'llanilish usullari keng yoritilgan. Shuningdek, berilgan ma'lumotlarni tozalash, tasvirlash va statistik ko'rsatkichlarni hisoblash bosqichlari batafsil ko'rib chiqilgan. Maqolada turli sohalarda taksonomiya usullarini qo'llashning ahamiyati va samaradorligi misollar orqali ko'rsatib berilgan.

Kalit so'zlar: Taksonomiya, ma'lumotlarni tahlil qilish, klasterlash, hierarxik klasterlash, K-means, statistik tahlil, dimensionalitetni qisqartirish.

Annotation: Taxonomy is an essential tool for systematization in science. It is widely used in various fields, from biological species to information technology, economics, pedagogy, and many more. This article analyzes the taxonomy methods and the initial analysis process of given data. The concept of taxonomy, its history, and its application in various fields are thoroughly covered. Additionally, the stages of cleaning, visualizing, and calculating statistical indicators of given data are examined in detail. The article demonstrates the importance and effectiveness of applying taxonomy methods in different fields through examples.

Keywords: taxonomy, data analysis, clustering, hierarchical clustering, K-means, statistical analysis, dimensionality reduction.

Аннотация: Таксономия является важным инструментом систематизации в науке. Она широко применяется в различных областях, начиная с биологических видов и заканчивая информационными технологиями, экономикой, педагогикой и многими другими сферами. В данной статье рассматриваются методы таксономии и процесс начального анализа данных. Подробно освещены понятие таксономии, её история и использование в различных областях. Также подробно рассматриваются этапы очистки,

визуализации и расчета статистических показателей данных. В статье демонстрируется важность и эффективность применения методов таксономии в различных областях на примерах.

Ключевые слова: таксономия, анализ данных, кластеризация, иерархическая кластеризация, K-means, статистический анализ, сокращение размерности

Taksonomiya tushunchasi va uning tarixi

Taksonomiya (lot. "taxo" – tartib, "tome" – bo'lim) bu obyektlar yoki hodisalarni muayyan mezonlar asosida guruhlash va tasniflash ilmidir. Taksonomiya qadim zamonlardan beri qo'llanilib keladi. Yunon olimi Aristotel hayvonlarni birinchi bo'lib tizimli tasniflagan. Zamonaviy taksonomiyaning asoschisi esa Karl Linney bo'lib, u 18-asrda tur va jinslarni tizimlashtirish bo'yicha ishlar olib borgan. Taksonomiya — ma'lumotlarni turkumlash va tasniflash jarayonidir. Sun'iy intellektda taksonomiya muhim ahamiyatga ega bo'lib, mashinaviy o'qitish modellari uchun ma'lumotlarni tuzilmaga solishda yordam beradi. Masalan, berilgan ma'lumotlar to'plamini turli kategoriya va sinflarga ajratish orqali SI modellari aniq prognoz va qarorlar qabul qiladi.

Taksonomiya asosiy bosqichlari

Taksonomiyani amalga oshirish quyidagi asosiy bosqichlardan iborat:

1. **Ob'yektlarning xususiyatlarini aniqlash:** Taksonomiya uchun birinchi bosqich ob'jekt yoki hodisalar haqida barcha kerakli ma'lumotlarni to'plashdan iborat.
2. **Olingan ma'lumotlarni guruhlash:** Obyektlar yoki hodisalar o'xshash belgilari asosida guruhlanadi.
3. **Tizimlashgan tuzilmani yaratish:** Guruhlangan ma'lumotlar ma'lum bir hierarxiya yoki ketma-ketlikda joylashtiriladi.

Taksonomiya usullari

Taksonomiya usullari tadqiqotlar va amaliy ishlarda turli xil usullarda qo'llaniladi. Quyida eng keng tarqalgan taksonomiya usullari keltirilgan:

1. Hierarxik klasterlash

Bu usulda ma'lumotlar yuqori darajadan pastga qarab, ketma-ket hierarxik guruhlariga bo'linadi. Hierarxik klasterlashning ikki turi mavjud: aglomerativ (pastdan yuqoriga) va diviziv (yuqoridan pastga). Amaliyotda bu usul biologiya, farmatsevtika va genetik tadqiqotlarda keng qo'llaniladi.

2. K-means klasterlash

Ma'lumotlarni belgilangan klasterlar soniga ajratish uchun ishlatiladi. K-means klasterlashda har bir klaster markaziy nuqtaga ega bo'lib, ma'lumotlar unga eng yaqin

klasterga biriktiriladi. Bu usul biznes tahlillarida, mijoz segmentatsiyasida va marketing strategiyalarida samarali ishlatiladi.

3. Dimensionalitetni qisqartirish usullari

Ma'lumotlar tahlilida ko'p o'lchovli ma'lumotlarni soddalashtirish uchun qo'llaniladi. **Principal Component Analysis (PCA)** va **t-SNE** kabi usullar asosiy ma'lumotlarni yo'qotmagan holda ma'lumotlar o'lchamlarini qisqartirish imkonini beradi. Bu usullar katta hajmdagi ma'lumotlar to'plamini boshqarish uchun muhim ahamiyatga ega.

Tajriba berilganlarni boshlang'ich tahlili

Ma'lumotlarni boshlang'ich tahlil qilish muhim bosqichlardan biri hisoblanadi. Bu bosqichda quyidagi jarayonlar amalga oshiriladi:

1. **Ma'lumotlarni tozalash:** Ma'lumotlar to'plamidagi noto'g'ri yoki to'ldirilmagan ma'lumotlarni bartaraf etish. Tozalash jarayoni yuqori sifatli va aniq tahlillar uchun zarur.
2. **Ma'lumotlarni tasvirlash:** Grafiklar, histogrammalar va chiziqli diagrammalar yordamida ma'lumotlarning umumiy ko'rinishini baholash. Vizualizatsiya tahlilchilarga ma'lumotlarni tezkor baholash imkonini beradi.
3. **Statistik ko'rsatkichlarni hisoblash:** O'rtacha qiymat, median, mod, dispersiya, standart og'ish kabi ko'rsatkichlarni hisoblash orqali ma'lumotlarning asosiy xususiyatlari aniqlanadi.
4. **Ma'lumotlarning tarqalishini o'rganish:** Klasterlash va tasniflash imkoniyatlarini o'rganish. Ma'lumotlarning tarqalishini bilish ma'lumotlarning murakkab tuzilmasini yaxshiroq tushunishga yordam beradi.

Taksonomiya va tahlilning amaliy qo'llanilishi Taksonomiya va ma'lumotlarni boshlang'ich tahlil qilish usullari quyidagi sohalarda keng qo'llaniladi:

1. **Biologiya:** Turlarning tasniflanishi, genetik tadqiqotlar. Biologik tur xilma-xilligini o'rganishda taksonomiya muhim o'rin tutadi.
2. **Iqtisodiyot:** Bozor segmentatsiyasi, mijoz xatti-harakatlarini tahlil qilish. Iqtisodiy model va prognoz yaratishda ma'lumotlarni guruhlash usullari ishlatiladi.
3. **Axborot texnologiyalari:** Sun'iy intellekt va tabiiy tilda ishlov berishda matnlarni tasniflash. Ma'lumotlarni indekslash va qidiruv tizimlarini yaratishda taksonomiya usullari muhim ahamiyatga ega.
4. **Sotsiologiya va pedagogika:** Ijtimoiy guruhlarni tahlil qilish. Ta'lim jarayonlarida talabalarni bilim darajasi bo'yicha tasniflash.

Xulosa

Taksonomiya va ma'lumotlarni boshlang'ich tahlil qilish usullari nafaqat ilmiy tadqiqotlarda, balki turli amaliy sohalarda ham muhim ahamiyatga ega. Ularni samarali qo'llash ma'lumotlardan chuqur va aniq xulosalar chiqarish imkonini beradi.

Maqolada keltirilgan taksonomiya usullari va ma'lumotlarni boshlang'ich tahlil qilish jarayonlari ilmiy izlanishlar va amaliyotda keng qo'llaniladi. Ushbu usullar yordamida katta hajmdagi ma'lumotlarni samarali boshqarish va ular asosida to'g'ri xulosalar chiqarish mumkin. Taksonomiya va tahlil jarayonining to'g'ri tashkil etilishi ilm-fan, biznes va ta'lim sohalarida samaradorlikni oshiradi. Bundan tashqari, zamonaviy klasterlash algoritmlarini qo'llash orqali yangi natijalarni olish va innovatsion yondashuvlarni ishlab chiqish imkoniyati oshadi. Taksonomiya va ma'lumotlar tahlili zamonaviy texnologik rivojlanishda muhim o'rin tutib, har bir sohada ma'lumotlarga asoslangan qarorlar qabul qilish jarayonini yengillashtiradi. Shu sababli, ushbu usullarni o'rganish va amaliyotga joriy etish orqali samaradorlikni yanada oshirish mumkin.

Foydalanilgan adabiyotlar

1. Linney, K. (1758). "Systema Naturae".
2. Everitt, B. S., & Dunn, G. (2001). "Applied Multivariate Data Analysis".
3. Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). "Data Clustering: A Review".
4. Hastie, T., Tibshirani, R., & Friedman, J. (2009). "The Elements of Statistical Learning".
5. Bishop, C. M. (2006). "Pattern Recognition and Machine Learning".
6. Qodirov, Sh. (2018). "Ma'lumotlarni tahlil qilish va tasniflash asoslari".
7. Karimov, B. (2020). "Amaliy statistika va klasterlash usullari".