

QAROR DARAXTLARI YORDAMIDA BANK MIJOZLARINING CHIQIB KETISHINI BASHORAT QILISH

(Farg'ona davlat texnika universiteti)

681-23 guruh talabasi

Iqboljon G'aniyev Ibrohimjon o'g'li

Annotatsiya: Ushbu maqolada qaror daraxtlari (decision tree) algoritmidan foydalanib, bank mijozlarining xizmatdan chiqib ketish ehtimolini bashorat qilish masalasi ko'rib chiqiladi. Maqolada mijozlar ma'lumotlari asosida modelni o'qitish, bashorat natijalarini tahlil qilish va samaradorlikni baholash usullari bayon etiladi. Amaliy qismda Python dasturlash tili va scikit-learn kutubxonasi yordamida model qurilgan.

Kalit so'zlar: Qaror daraxti, bank mijozlari, mashinani o'rganish, bashorat, Python, scikit-learn, model, ma'lumotlar tahlili.

Annotation: In this article, the problem of predicting the likelihood of bank customers churning is examined using the decision tree algorithm. The article describes methods for training the model based on customer data, analyzing prediction results, and evaluating performance. In the practical section, a model is built using the Python programming language and the scikit-learn library.

Keywords: Decision tree, bank customers, machine learning, prediction, Python, scikit-learn, model, data analysis

Аннотация: В данной статье рассматривается задача прогнозирования вероятности оттока клиентов банка с использованием алгоритма решающих деревьев. Описаны методы обучения модели на основе данных о клиентах, анализ полученных результатов и оценка эффективности модели. В практической части представлено построение модели прогнозирования оттока с использованием языка программирования Python и библиотеки scikit-learn.

Ключевые слова: Решающее дерево, клиенты банка, машинное обучение, прогнозирование, Python, scikit-learn, модель, анализ данных

Kirish

Hozirgi kunda bank sektorida mijozlarni saqlab qolish muhim strategik masalalardan biridir. Mijozlarning xizmatdan chiqib ketish sabablari turlicha bo'lishi mumkin: xizmat sifati, kredit shartlari, raqobatchilar takliflari va boshqalar. Shu sababli, banklar sun'iy intellekt (AI) va mashinani o'rganish (Machine Learning) usullaridan foydalangan holda mijozlarning chiqib ketish ehtimolini oldindan aniqlashga intilmoqda.

Nazariy qism: Qaror daraxtlari (Decision Tree)

Qaror daraxti — bu mashinaviy o'qitish (Machine Learning) algoritmi bo'lib, ma'lumotlarni “**Agar, unda...**” shaklidagi qoidalar asosida tahlil qiladi va ularni turkumlaydi yoki bashorat qiladi. Ushbu algoritmlar qarorlarni daraxt ko'rinishida modellashtiradi: ildiz tugundan (asosiy belgi) boshlab, ichki tugunlarda ma'lumot turli shartlarga ajratiladi va oxirgi barg tugunlarda yakuniy natija (bashorat) hosil bo'ladi.

Qaror daraxtlari **tasniflash (classification)** va **regressiya (regression)** muammolarini hal qilishda ishlatiladi. Ular, ayniqsa, mijozlarning xatti-harakatlarini, risk darajasini yoki chiqib ketish ehtimolini aniqlashda keng qo'llaniladi.

Qaror daraxtining asosiy elementlari

Ildiz tugun (Root Node) — daraxtning boshlanish nuqtasi, ya'ni asosiy belgi bo'yicha ajratish shu tugundan boshlanadi.

Ichki tugun (Internal Node) — qaror shartini ifodalaydi va ma'lumotlarni bo'linishga yo'naltiradi.

Barg tugun (Leaf Node) — yakuniy natijani (bashoratni) ko'rsatadi.

Shoxlar (Branches) — tugunlarni bir-biriga bog'lovchi yo'llar bo'lib, ular shartlarning natijasini ifodalaydi

Qaror daraxtlarining afzalliklar

Tushunarli va vizual — qaror daraxti inson uchun oson talqin qilinadi.

Kichik ma'lumotlar bilan ham ishlay oladi — katta ma'lumotlar bazasi talab qilmaydi.

Kategorik va sonli ma'lumotlarga mos — har ikki turdagi ma'lumotlarni qayta ishlay oladi.

Izohlanadigan model — qarorlar qanday qabul qilinganini kuzatish mumkin.

Kamchiliklari

Juda ko'p shartlar bo'lganda daraxt **murakkablashadi** (overfitting holati yuz beradi).

Kichik o'zgarishlar natijaga sezilarli ta'sir qilishi mumkin.

Ba'zida **bashorat aniqligi** ansambl modellarga nisbatan pastroq bo'ladi.

Qaror daraxtining ishlash prinsipi

Ma'lumot to'plami asosida har bir atribut (belgi) uchun **axborot foydasi** (Information Gain) yoki **Gini indeksi** hisoblanadi.

Eng katta foyda keltiruvchi atribut ildiz tugun sifatida tanlanadi.

Ma'lumot shu atribut bo'yicha kichik bo'laklarga ajratiladi.

Har bir bo'lak uchun shu jarayon takrorlanadi, to ularning toifalari aniq bo'lmaguncha yoki chuqurlik chegarasiga yetguncha.

Bank sohasida qo'llanilishi

Bank tizimlarida qaror daraxtlari **mijozlarning chiqib ketishini (churn prediction)** oldindan bashorat qilishda keng qo'llaniladi. Masalan, model quyidagicha ishlaydi:

Kiruvchi ma'lumotlar: **mijoz yoshi, balans miqdori, kredit balli.**

Model har bir belgi asosida qaror qabul qiladi:

“Agar mijozning balansi 15 000 dan kam va kredit balli 550 dan past bo‘lsa chiqib ketish ehtimoli yuqori.”

“Agar mijozning balansi 40 000 dan yuqori bo‘lsa → bankda qolish ehtimoli katta.”

Shunday qilib, qaror daraxti bankka mijozlarni yo‘qotmaslik uchun **oldindan ogohlantiruvchi tahlil** olib borish imkonini beradi.

Amaliy qism (Python kodi)

```
import pandas as pd
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score, classification_report

data = {
    "CreditScore": [600, 750, 500, 720, 650, 580, 790, 400],
    "Age": [40, 35, 50, 42, 30, 28, 60, 45],
    "Balance": [60000, 80000, 30000, 90000, 15000, 20000, 95000, 12000],
    "IsActiveMember": [1, 0, 1, 1, 0, 1, 0, 0],
    "Gender": ["Male", "Female", "Male", "Female", "Male", "Female", "Male",
    "Female"],
    "Exited": [0, 1, 0, 1, 0, 0, 1, 0]
}
df = pd.DataFrame(data)
label = LabelEncoder()
df["Gender"] = label.fit_transform(df["Gender"])
X = df.drop("Exited", axis=1)
Y = df["Exited"]
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.3,
random_state=42)
model = DecisionTreeClassifier(max_depth=5, random_state=42)
model.fit(X_train, Y_train)
Y_pred = model.predict(X_test)
print(accuracy_score(Y_test, Y_pred))
print(classification_report(Y_test, Y_pred))
```

Natijalar

Model bank mijozlarining yosh, balans va kreditlar soni kabi omillarga asoslanib, chiqib ketish ehtimolini aniqlay oldi. Qaror daraxti orqali model mijozlarni ikki toifaga ajratadi: qoladiganlar va chiqib ketadiganlar.

Xulosa

Qaror daraxtlari yordamida bank mijozlarining chiqib ketishini bashorat qilish natijasi shuni ko'rsatadiki, bu usul oddiy, tushunarli va samarali hisoblanadi. Kelgusida modelga ko'proq atributlar (mijozning daromadi, xizmat muddati, kredit tarixlari va h.k.) qo'shish orqali aniqlikni yanada oshirish mumkin.

Foydalanilgan adabiyotlar

1. **Mitchell, T. M.** *Machine Learning*. McGraw-Hill, 1997.
— Mashinani o'rganish algoritmlari, jumladan qaror daraxtlarining nazariy asoslari yoritilgan.
2. **Han, J., Kamber, M., Pei, J.** *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2012.
— Ma'lumotlarni tahlil qilish va tasniflash algoritmlarining amaliy va nazariy jihatlari bayon etilgan.
3. **James, G., Witten, D., Hastie, T., Tibshirani, R.** *An Introduction to Statistical Learning*. Springer, 2013.
— Klassifikatsiya va regressiya modellariga, jumladan decision tree algoritmiga kirish.
4. **Pedregosa, F. et al.** "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research*, 12, 2011, pp. 2825–2830.
— Python tilidagi scikit-learn kutubxonasining imkoniyatlari va algoritmlari haqida.
5. **Brownlee, J.** *Machine Learning Mastery With Python*. Machine Learning Mastery, 2016.
— Python yordamida mashinani o'rganish modellarini qurish bo'yicha amaliy qo'llanma.