

**GRAF NEYRON TARMOQLARINI TASNIFLASH MASALALARINI YECHISHGA
QO'LLASH**

Boltibayev Shuxratjon Komiljanovich

Namangan davlat universiteti dotsenti

Email: sh.boltibayev@gmail.com

Rahimov Boburjon Ma'rufjon o'g'li

Toshkent Kimyo Xalqaro Universiteti Namangan filiali magistranti

DOI: <https://doi.org/10.5281/zenodo.20215051>

Annotatsiya. Obyektlarni tasniflash muammolari sanoatda juda keng tarqalgan. Ko'pgina hollarda, ular grafik ma'lumotlarini hisobga olmasdan, asosiy xususiyatlar (klassik mashina o'rganish algoritmlari) yordamida hal qilinadi. Tadqiqotning gipotezasi shundan iborat ediki, grafli tuzilishini hisobga oladigan usullar yuqori sifatli modellarni beradi.

Ushbu maqolada grafli ma'lumotlaridan foydalangan holda tasniflash muammosini hal qilishning bir nechta yondashuvlari taqdim etilgan va sinovdan o'tkazilgan. Ular telekommunikatsiya kompaniyasining ma'lumotlarini qayta ishlashda qo'llanilishi mumkin.

Kalit so'zlar: neyron tarmoqlari, grafik ma'lumotlari, grafik cho'qqilari ma'lumotlari.

**ПРИМЕНЕНИЕ ГРАФОВЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ
КЛАССИФИКАЦИИ**

Болтибаев Шухратжон Комилжанович

Доцент, Наманганский государственный университет

Рахимов Бобуржон Маъруфжон огли

Магистрант Наманганского филиала Международного университета Кимё в Ташкенте

Аннотация. Задачи классификации объектов очень распространены в промышленности. В большинстве случаев они решаются с использованием базовых признаков (классических алгоритмов машинного обучения) без учета графовых данных. Гипотеза

исследования заключалась в том, что методы, учитывающие структуру графа, обеспечивают высококачественные модели.

В данной статье представлены несколько подходов к решению задачи классификации с использованием графовых данных. Они могут быть применены в обработке данных телекоммуникационной компании.

Ключевые слова: нейронные сети, графовая информация, вершин графа.

APPLICATION OF GRAPH NEURAL NETWORKS TO SOLVING CLASSIFICATION PROBLEMS

Boltibayev Shuhratjon Komiljanovich

Associate Professor, Namangan State University

Rahimov Boburjon Ma'rufjon o'g'li

Master's student, Namangan branch of Kimyo international university in Tashkent

Abstract. Object classification problems are very common in industry. In most cases, they are solved using basic features (classical machine learning algorithms) without taking graph data into account. The research hypothesis was that methods that take graph structure into account produce high-quality models.

This article presents several approaches to solving classification problems using graph data. They can be applied to data processing at a telecommunications company.

Keywords: neural networks, graph information, graph vertices

Neyron tarmoqlari (ayniqsa, o'rama neyron tarmoqlari) tasvir, signal va matnni qayta ishlashda katta bum yaratdi. Biroq, ularni to'g'ridan-to'g'ri grafiklarga qo'llash mumkin emas - grafiklar, yuqorida tavsiflangan ob'ektlardan farqli o'laroq, aniq boshlang'ich nuqtaga yoki tartibga ega emas va odatda dinamikdir. Grafik neyron tarmoqlari va undan ilg'or yondashuvlar - grafik konvolyutsion neyron tarmoqlari (GCN) - bu muammoni hal qilishi mumkin. Ular tepaliklar to'plamlaridan ma'lumotlarni to'playdi, bu grafik tuzilishini aniq tasvirlash imkonini beradi va shuningdek, grafik tepaliklarining xususiyatlari bilan ishlashga qodir.

Masalani quyidagicha shakllantiramiz:

G grafi berilgan, unda:

- V tugunlar to'plami;
- A qo'shnilik matritsasi (biz ikkilik deb faraz qilamiz);
- X tugunlar xarakteristikalar matritsasi.

GCN usuli to'g'ridan-to'g'ri grafik tuzilishiga va uning "qo'shnilari" xususiyatlariga tayanadi. Ushbu yondashuvning asosiy g'oyasi grafik uchlari uchun bir xil "konvolyutsiya" dan foydalanishdir. Biroq, tasvirlardan farqli o'laroq, grafli neyron tarmoqlarida grafning o'zi neyron tarmog'ining tuzilishini belgilaydi. Uni joylashtirish ko'p bosqichli agregatsiya va atrofdagi uchlarning xususiyatlarini o'zgartirishga asoslangan bo'ladi. Asosiy savol bu agregatlarni qanday qurish, ya'ni grafik orqali ma'lumotni qanday tarqatishdir. Keling, William L. Hamilton, Rex Ying va Jure Leskovec tomonidan [1] da taklif qilingan g'oyani ko'rib chiqaylik.

Eng asosiy yondashuv operatorga kiruvchi vektorlarni o'rtacha hisoblash va natijaga neyron tarmog'ini qo'llashdir. Keyin, yuqorida tavsiflanganidek,

$$\begin{cases} h_v^0 = x_v \\ h_v^k = \sigma \left(W_k \sum_{u \in N(v)} \frac{h_u^{k-1}}{|N(v)|} + B_k h_v^{k-1} \right), \forall k \in 1, \dots, K, \\ z_v = h_v^k \end{cases}$$

bunda h_v^i — i -qatlamga joylashtirishning tasviri;

$\sigma(x)$ — ba'zi chiziqli bo'lmagan transformatsiyalar;

x_v — v tugunning xususiyatlari vektori;

z_v — v tugunning oxirgi vektor ko'rinishi;

W_k, B_k — tarmoq parametrlari.

Tarmoq parametrlarini tanlash uchun biz yo'qotish funksiyasini aniqlashimiz va stoxastik gradient tushishini bajarishimiz kerak.

Ikkita variant mavjud:

- Unsupervised Learning (nazoratsiz o'rganish). Bu holda biz tugun vektorlarini qurish muammosini hal qilishimiz mumkin. Bu holda, node2vec usulidagi kabi bir xil metrikadan metrika sifatida foydalanish mumkin;

— Supervised Learning (nazoratli o'rganish). Bunday muammoning misoli tugunlar sinfini aniqlashdir. Bu holda biz ma'lum tugunlar qaysi sinfga tegishli ekanligini bilamiz. Keyin, grafli ma'lumotlarga asoslanib (GCN usullaridan foydalangan holda), biz ulardan qaysi biri boshqalarga tegishli ekanligini taxmin qilishimiz mumkin. Masalan, ushbu usuldan foydalanib, biz ma'lum bir shaxs telekommunikatsiya kompaniyasining ma'lumotlari asosida ma'lum bir jamoaning muxlisi ekanligini aniqlash mumkin.

Ikkiлик tasniflash muammosini hal qilish uchun biz o'zaro entropiya yo'qotishini masalaga moslashtiramiz. Keyin yo'qotish funksiyasi quyidagicha ko'rinishda bo'ladi:

$$L = \sum_{v \in V} y_v \log\left(\sigma\left(z_v^T \theta\right)\right) + (1 - y_v) \log\left(1 - \sigma\left(z_v^T \theta\right)\right),$$

bunda z_v — v tugunlar vektori;

θ — tasniflash uchun og'irliklar vektori;

$y_v \in \{0, 1\}$ — haqiqiy tugunlar sinfi.

Ushbu yondashuvning afzalligi shundaki, u o'rganish (grafning bir qismida W_k va B_k parametrlarini tanlash) imkonini beradi. Keyin, ushbu parametrlarga asoslanib, biz butun modelni qayta hisoblamasdan yangi tugunlar uchun bashoratlarni hisoblashimiz mumkin.

GCN eng asosiy arxitektura bo'lib, ko'pincha amalda qo'llaniladi. Biroq, quyidagi yondashuvlar ham keng tarqalgan va qo'llanilmoqda:

— Grafli neyron tarmoqlari (GNN) — o'ramalarsiz klassik grafli neyron tarmoqlari;

— GraphSAGE — kengroq mumkin bo'lgan agregatsiya funktsiyalariga ega GCN g'oyalarning kengaytmasi;

— Grafik diqqat tarmoqlari (GAT) — agregatsiya og'irliklarini avtomatik (alohida tarmoq yordamida) sozlash imkonini beruvchi diqqat mexanizmiga ega grafli neyron tarmoqlari.

Ushbu tadqiqot doirasida modellarni quyidagi loyihada sinab ko'rish mumkin:

— Soxta obunachilarni aniqlash. Kompaniyaning asosiy vazifalaridan biri shubhali faoliyat ko'rsatadigan va noqonuniy xatti-harakatlarni sodir etadigan spammer obunachilarni aniqlash;

— Foydalanuvchilarning o'chib ketishini bashorat qilish. Bu vazifa mijozning provayderlarni almashtirishga moyilligini bashorat qilishni o'z ichiga oladi va asosan grafli tuzilishiga bog'liq (agar

ma'lum bir mijozning kontaktlari provayderlarni o'zgartirsa, ular ham shunday qilishlari ehtimoli yuqori);

— Obunachilarning jinsini bashorat qilish modeli. Bugungi haqiqatda marketing telekommunikatsiya kompaniyasi faoliyatining ajralmas qismi hisoblanadi. Foydalanuvchilarning demografik xususiyatlarini (jinsi, yoshi va boshqalar) aniqlash ko'pincha zarur. Bu parametrlar raqamni ro'yxatdan o'tkazishda ko'rsatiladi, lekin aslida SIM-kartalar ko'pincha ular ro'yxatdan o'tganlardan boshqa odamlar tomonidan ishlatiladi, bu esa bu vazifani juda dolzarb qiladi.

Foydalanilgan adabiyotlar:

1. William L. Hamilton, Rex Ying, Jure Leskovec: Representation Learning on Graphs: Methods and Applications: <https://arxiv.org/pdf/1709.05584.pdf>
2. Yoav Goldberg and Omer Levy: word2vec Explained: Deriving Mikolov et al.'s Negative-Sampling Word-Embedding Method <https://arxiv.org/pdf/1402.3722.pdf>
3. S. Cao, W. Lu, and Q. Xu. Deep neural networks for learning graph representations. In AAAI, 2016