

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»
ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ СИСТЕМНОГО АНАЛІЗУ

Дипломна робота на тему:
«Методи прогнозування рекламної активності на
телебаченні України»

Виконала:

Клімова Олена Віталіївна, студентка
групи КА-43

Науковий керівник:

професор, д.т.н. Зайченко О.Ю.

Вступ

- ▶ **Об'єкт дослідження:**

Реклама на телебаченні України

- ▶ **Предмет дослідження:**

Методи прогнозування рекламної активності брендів за допомогою МГУА та нейронної мережі Back Propagation

- ▶ **Мета дослідження:**

Створення системи для отримання прогнозу рекламної активності у вигляді рейтингів



Постановка задачі

- ▶ Провести дослідження рекламного ринку, визначитися з даними та методами прогнозування.
- ▶ Реалізувати обрані алгоритми прогнозування та представити результати їх роботи у вигляді графіків та обраних показників.
- ▶ Порівняти результати прогнозованих рейтингів різними методами з реальними показниками.
- ▶ Проаналізувати отримані результати, визначити та описати основні переваги, недоліки та особливості досліджуваних моделей в контексті їх застосування в задачах прогнозування.



Постановка задачі

- ▶ Дане дослідження та відповідний програмний продукт розроблені для медійного агентства ZenithOptimedia, яке працює на українському ринку з 1998 року. Воно входить в рекламно-комунікаційний холдинг Publicis Groupe та має філіали в 74 країнах світу. Сьогодні це четвертий за величиною рекламно-комунікаційний холдинг світу і перший на європейському ринку.
- ▶ ZenithOptimedia надає послуги зі стратегічного і тактичного медіапланування, медіаконсалтінгу і закупівель реклами. Протягом всієї кампанії фахівці постійно аналізують активності за допомогою використання трекінгових інструментів і вносять необхідні зміни для підвищення ефективності інвестицій.



Актуальність роботи

- ▶ При створенні рекламної кампанії медіапленер має орієнтуватися на активність конкурентів, щоб побудувати свою стратегію в залежності від них, спираючись на свій бюджет.
- ▶ Представлення результатів моделювання клієнтам для їх аналізу у відділі продажів.



Медіапланування

- ▶ Медіапланування – складання плану рекламної кампанії при оптимальному виборі необхідних каналів розміщення реклами в ЗМІ на основі даних маркетингових і медіадосліджень.
- ▶ У медіаплані рекламної кампанії описуються обрані типи реклами, терміни розміщення, вартість (пакетна або разова покупка, спонсорство і т.п.), а також ефективність проведеної рекламної кампанії.



Вибірка для проведення дослідження

	Brand	Segment	Group	Type Spot	Prime/Offprime	Year	Length	Week	rat%	atv	EQGrp
0	Carte Noire	Super Premium	кофе растворимый	0	0	2014	31	14	15.65375	225.0	2.31
1	Carte Noire	Super Premium	кофе растворимый	0	1	2014	29	14	15.65375	225.0	1.55
2	Carte Noire	Super Premium	кофе растворимый	0	1	2014	29	15	16.00159	230.0	0.91
3	Carte Noire	Super Premium	кофе растворимый	0	1	2014	29	18	15.35007	221.0	0.05
4	Carte Noire	Super Premium	кофе растворимый	0	0	2014	31	17	14.61458	210.0	0.06

Дана вибірка заснована на даних рекламної активності по категорії «кава». Джерелом даних є GFK, самі дані отримані за допомогою програмного забезпечення Markdata. Обсяг вибірки складає 37752 рядків, в той час як часовими рамками є період з 2014 по 2018 рік.



Показники вибірки

- ▶ 1. Brand – назва бренду.
- ▶ 2. Segment – сегмент.
- ▶ 3. Group – група товарів.
- ▶ 4. Type Spot – вид рекламної активності.
- ▶ 5. Prime/Offprime – час виходу ролику в ефір.
- ▶ 6. Year – рік виходу ролику.
- ▶ 7. Length – довжина ролику.
- ▶ 8. Week – тиждень виходу ролику.
- ▶ 9. Rat% – рейтинг телеперегляду на даному тижні.
- ▶ 10. Atv (Average audience rate) – середня кількість рейтингів для обраної цільової аудиторії.
- ▶ 11. EQGrp – вихідна величина, зважені рейтинги, що оцінюють ефективність донесення рекламного повідомлення.



Блок-схема процесу



Метод групового урахування аргументів

- ▶ Будуємо часткові описи:

$$y_s = \varphi(x_i, x_j) = a_0 + a_i x_i + a_j x_j, \quad s = 1 \dots C_n^2 \text{ (лінійні)}$$

- ▶ Визначаємо коефіцієнти цих моделей за МНК.

$$\overline{a_0}, \overline{a_1}, \dots, \overline{a_j}, \dots, \overline{a_N}, \overline{a_{11}}, \dots, \overline{a_{ij}}, \dots, \overline{a_{NN}}.$$

- ▶ Шукаємо оцінку за критерієм регулярності:

$$\overline{\delta_s^2} = \frac{1}{N_{\text{перев}}} \sum_{i=1}^{N_{\text{перев}}} [Y(k) - \overline{Y}_s(k)]^2$$

- ▶ Визначаємо F найкращих моделей. Обрані моделі подаємо на другий ряд.

$$z_1 = \varphi(y_i, y_j) = a_0^{(2)} + a_1^{(2)} y_i + a_2^{(2)} y_j + a_3^{(2)} y_i^2 + a_4 y_i y_j + a_5 y_j^2$$

- ▶ Процес побудови рядів відбувається до тих пір, поки середній квадрат помилки буде зменшуватися. Коли на шарі m помічаємо збільшення помилки, процес синтезу моделі припиняємо.



Метод зворотного поширення помилки

- ▶ Ініціалізація вагів.
- ▶ Вхідний нейрон відправляє сигнал нейронам прихованого шару.
- ▶ Прихований нейрон сумує зважені вхідні сигнали та застосовує функцію активації. Посилає результат всім елементам вихідного шару.

$$z_in_j = v_{0j} + \sum_i^n x_i * v_{ij} \quad z_j = f(z_in_j).$$

- ▶ Вихідний нейрон сумує зважені вхідні сигнали та рахує вихідний сигнал.

$$y_in_k = w_{0k} + \sum_j z_j * w_{jk} \quad y_k = f(y_in_k).$$

- ▶ Вихідний нейрон рахує помилку та зміну вагів.

$$\sigma_k = (t_k - y_k) * f'(y_in_k) \quad \Delta w_{jk} = \alpha * \sigma_k * z_j$$

- ▶ Прихований нейрон сумує вхідні помилки, рахує помилку та зміну вагів.

$$\sigma_in_j = \sum_{k \in 1}^n \sigma_k * w_{jk} \quad \sigma_j = \sigma_in_j * f'(z_in_j) \quad \Delta v_{ij} = \alpha * \sigma_j * x_i$$

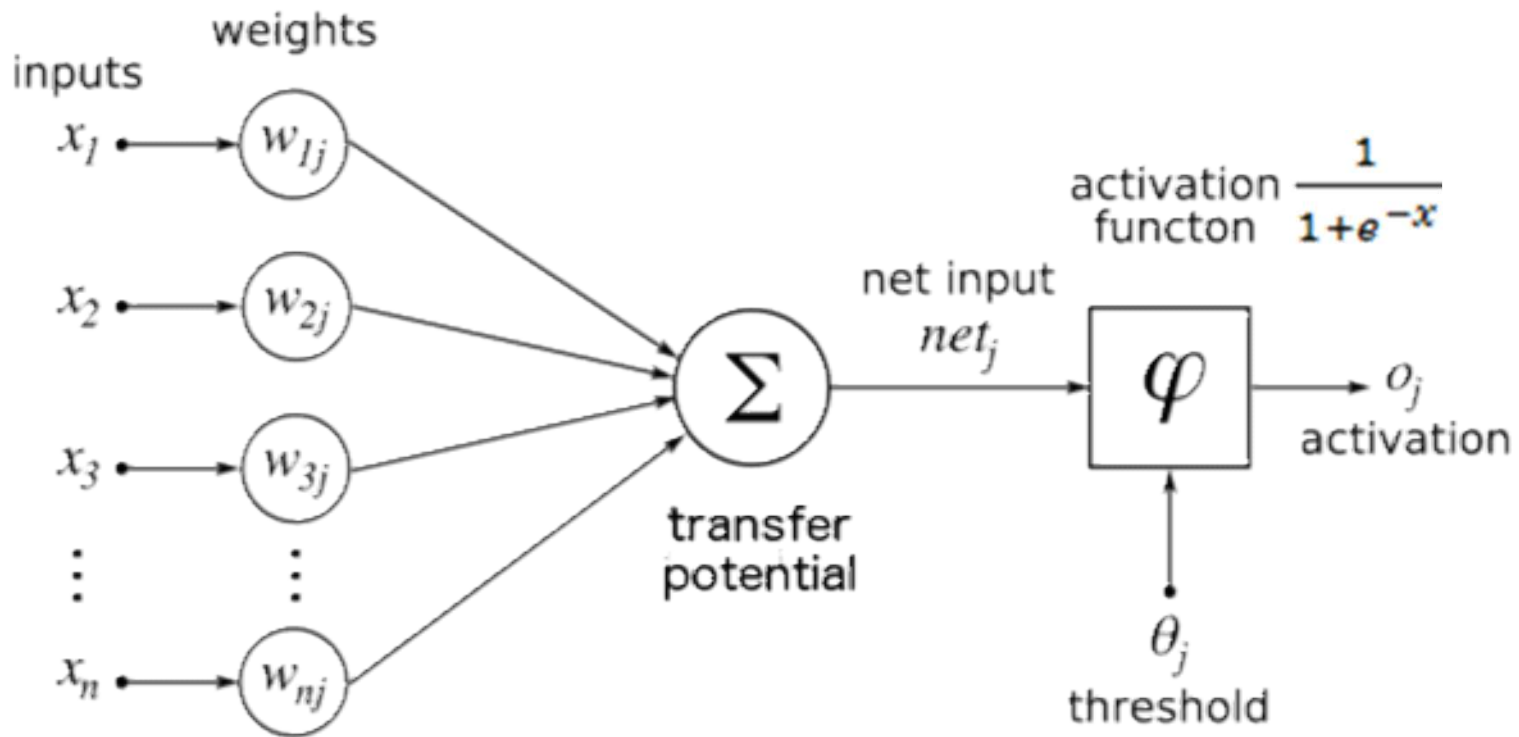
- ▶ Вихідні та приховані нейрони змінюють ваги.

$$w_{jk}(new) = w_{jk}(old) + \Delta w_{jk} \quad v_{ij}(new) = v_{ij}(old) + \Delta v_{ij}$$

- ▶ Перевірка умови закінчення роботи алгоритму.
-



Метод зворотного поширення помилки



Метрики

- ▶ Коефіцієнт детермінації


$$R^2 = \frac{\text{var}(\hat{y})}{\text{var}(y)}$$

- ▶ Середньоквадратична похибка

$$MSE = \frac{1}{n} \sum_{k=1}^n (y(k) - \hat{y}(k))^2$$

- ▶ Середня абсолютна похибка

$$MAE = \frac{1}{n} \sum_{k=1}^n |y(k) - \hat{y}(k)|$$



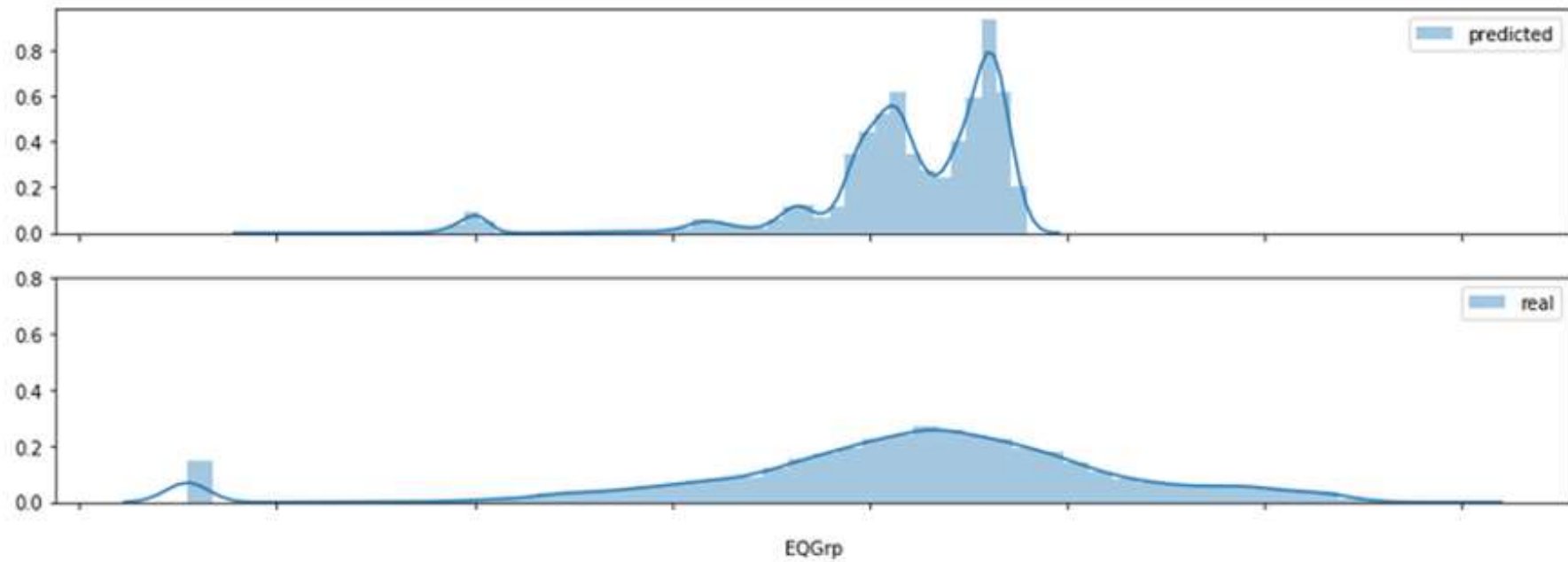
Вибір кращої моделі

Модель	R ²	MSE	MAE
МГУА	0.242495	3.755383	1.376830
НМ	0.501076	2.511851	1.003958



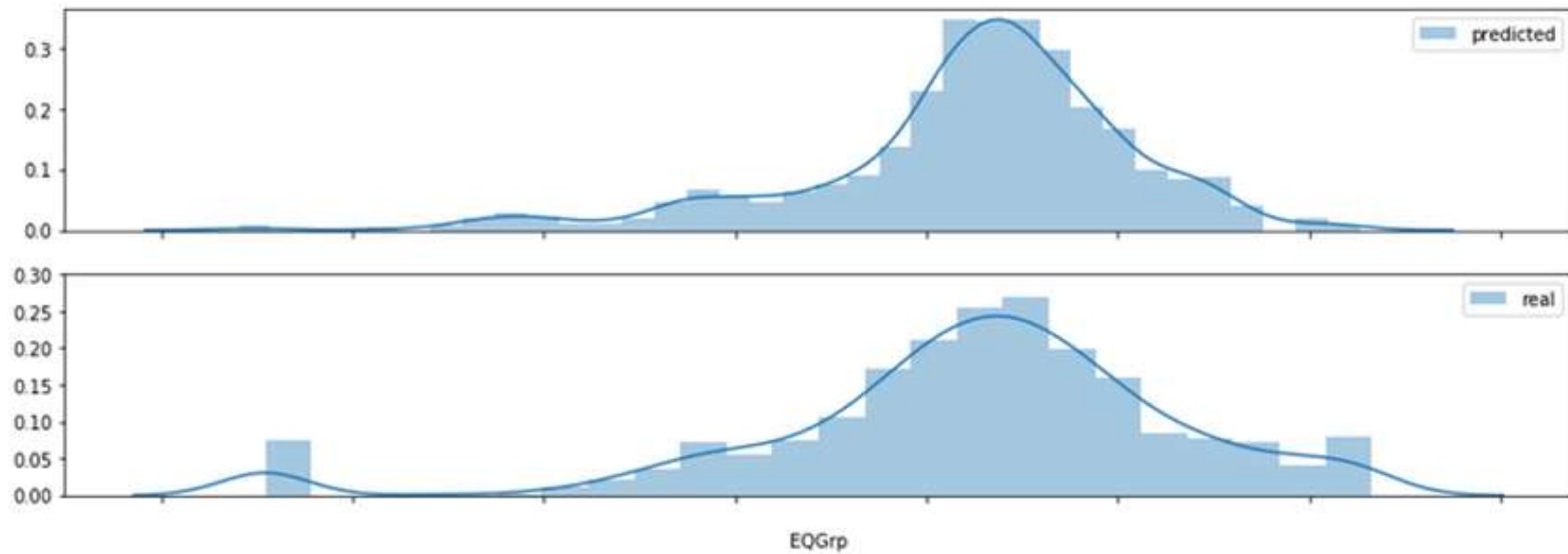
Аналіз результатів

МГУА



Аналіз результатів

Обернене поширення помилки



Висновки

- ▶ Проведено детальний аналіз предметної області та виявлено найбільш впливові характеристики рекламної активності на телебаченні.
- ▶ Сформовано вибірку, на основі якої побудовано прогноз за допомогою МГУА та нейронної мережі Back Propagation.
- ▶ Проведено аналіз результатів, який показав, що нейронна мережа спрацювала краще.
- ▶ Тези даної роботи були подані на міжнародну конференцію SAIT 2018.



Шляхи подальшого розвитку

- ▶ Використання інших методів прогнозування та їх порівняльна характеристика.
- ▶ Вдосконалення моделей для досягнення більшої точності.
- ▶ Застосування отриманих результатів в роботі медіапенерів.



ДЯКУЮ ЗА УВАГУ!

