

УДК 551.578.48

DOI <https://doi.org/10.66941/3106-9037/2025/iss3pp28-36>**Темір Нұрмахан***Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан***БАНК ПЕН КЛИЕНТТЕР АРАСЫНДАҒЫ ӨЗАРА ӘРЕКЕТТЕР ПРОЦЕСІНДЕГІ ӘЛСІЗ  
НҮКТЕЛЕРДІ АНЫҚТАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМІН ҚОЛДАНУ**

**Темір Нұрмахан**, магистрант, Ақпараттық технологиялар факультеті, Жасанды интеллект және BIG DATA кафедрасы, магистрант, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.

**Аннотация.** Қаржы секторын цифрлық трансформациялау және міндеттемелер үшін бәсекелестіктің күшеюі жағдайында коммерциялық банктер өз маркетингтік коммуникацияларын оңтайландыру қажеттілігіне тап болады. Депозиттерді тартудың дәстүрлі моделі – жаппай телемаркетинг («суық қоңыраулар») – төмен конверсиялық көрсеткіштер мен жоғары операциялық шығындармен сипатталады, бұл бизнес-процестерде тосқауылдар тудырады. Бұл зерттеу машинамен оқыту алгоритмдеріне негізделген мұндай шектеулерді анықтау және жою әдістемесін ұсынады. Эмпирикалық деректерге (11 162 бақылау) сүйене отырып, екілік классификация модельдері (логистикалық регрессия, Random Forest, Gradient Boosting) құрылып, валидацияланды. Алдын ала аналитикаға көшу мақсатталмаған трафиктің 43%-ға дейінгі бөлігін анықтауға, клиент тарту шығынын (CAC) азайтуға және сату процесін өзгертуге мүмкіндік беретіні дәлелденді. Экономикалық тиімділік тұрғысынан Precision және Recall көрсеткіштерін түсіндіруге ерекше назар аударылады.

**Түйін сөздер:** банктік маркетинг, машинамен оқыту, екілік жіктеу, таршылықтар, болжамды аналитика, Градиенттік бустинг, Рэндэм Форест, бизнес-процестерді оңтайландыру, деректерді басқару.

**Темір Нұрмахан***Казахский национальный университет им. Аль-Фараби, Алматы, Казахстан***Использование алгоритма машинного обучения для выявления слабых мест в процессе  
взаимодействия банка и клиентов**

**Темир Нурмахан**, магистрант, факультет информационных технологий, кафедра искусственного интеллекта и BIG DATA, магистрант, Казахский национальный университет им. Аль-Фараби, Алматы, Казахстан.

**Аннотация.** В условиях цифровой трансформации финансового сектора и усиления конкуренции за обязательства коммерческие банки сталкиваются с необходимостью оптимизации своих маркетинговых коммуникаций. Традиционная модель привлечения депозитов-массовый телемаркетинг («холодные звонки») – характеризуется низкими показателями конверсии и высокими операционными затратами, что создает барьеры в бизнес-процессах. Это исследование предлагает методологию обнаружения и устранения таких ограничений на основе алгоритмов машинного обучения. На основе эмпирических данных (11 162 наблюдения) были созданы и проверены модели бинарной классификации (логистическая регрессия, Random Forest, Gradient Boosting). Доказано, что переход к предварительной аналитике позволяет выявить до 43% нецелевого трафика, снизить затраты на привлечение клиентов (CAC) и изменить процесс продаж. С точки зрения экономической эффективности особое внимание уделяется интерпретации показателей Precision и Recall.

**Ключевые слова:** банковский маркетинг, машинное обучение, бинарная классификация, узкие места, прогнозная аналитика, градиентный бустинг, Рэндэм Форест, оптимизация бизнес-процессов, управление данными.

**Temir Nurmakhan**

*Al-Farabi Kazakh National University, Almaty, Kazakhstan*

## **Using a machine learning algorithm to identify weak points in the process of interaction between the bank and customers**

**Temir Nurmakhan**, master's Student, Faculty of Information Technologies, Department of artificial intelligence and BIG DATA, Master's student, Al-Farabi Kazakh National University, Almaty, Kazakhstan.

**Abstract.** In the context of the digital transformation of the financial sector and increased competition for obligations, commercial banks are faced with the need to optimize their marketing communications. The traditional model of attracting deposits – mass telemarketing ("cold calls") – is characterized by low conversion rates and high operating costs, which creates blockages in business processes. This study provides a methodology for identifying and eliminating such constraints based on machine learning algorithms. Based on empirical data (11,162 observations), binary classification models (logistic regression, Random Forest, Gradient Boosting) were created and validated. It has been proven that the transition to pre-analytics allows you to identify up to 43% of non-targeted traffic, reduce customer engagement loss (CAC), and change the sales process. Particular attention is paid to the interpretation of Precision and Recall indicators in terms of economic efficiency.

**Keywords:** banking marketing, machine learning, binary classification, bottlenecks, predictive analytics, gradient busting, Randham Forest, Business Process Optimization, Data Management.

**1. Кіріспе.** Бүгінгі экономикада деректер қаржылық институттардың бәсекеге қабілеттілігін анықтайтын негізгі активке айналуға. Екінші деңгейлі банктер тұрақты қаржыландыруды қамтамасыз ету мақсатында бөлшек салым өнімдерін белсенді түрде дамытуда. Алайда осы өнімдерді сату арналары көбінесе консервативті болып қала береді. Телемаркетинг тұтынушылармен өзара әрекеттесудің негізгі құралы болып қала береді — контакт-орталық операторларының клиенттер базасына тікелей қоңырау шалуы [1].

Бұл тәсілдің мәселесі – оның төмен таңдамалылығы. Банктер мүмкіндігінше көп әлеуетті салымшыларға жету үшін «кілемдік жабу» стратегияларын қолданады. Бұл жүйелік мәселелерге әкеледі:

1. Төмен конверсиялық көрсеткіштер: Статистикаға сәйкес, банктік сектордағы суық қоңыраулардың тиімділігі сирек 1–3%-дан асады, бұл ресурстардың үлкен ысырапқа ұшырағанын көрсетеді.
2. Операциялық тосқауылдар: білікті қызметкерлердің уақыты мен телефон желілерінің сыйымдылығы бас тартуларды өңдеуге жұмсалады, ал кезектер немесе ресурстардың жетіспеушілігі салдарынан адал клиенттер назардан тыс қалуы мүмкін [3].
3. Беделдік тәуекелдер: шабуылшыл маркетинг адалдықты (NPS) төмендетіп, клиенттердің кетуіне (Churn Rate) ықпал етеді.

Зерттеудің өзектілігі: банк басқаруындағы парадигманы өзгерту қажеттілігімен байланысты: колл-орталық қызметкерлерін кең көлемде ұлғайтудан ақылды шешім қабылдауды қолдау жүйелерін интенсивті пайдалануға. Жұмыстың мақсаты – сату процесіндегі шектеулерді анықтауға және оларды тұтынушы мінез-құлқын болжамдық модельдеу арқылы оңтайландыруға арналған алгоритмдік тәсілді әзірлеу.

## **2. Материалдар мен әдістер**

### **2.1. Деректерді сипаттау және зерттеушілік деректер талдауы (EDA)**

Зерттеудің эмпирикалық негізін 11 162 маркетингтік байланыс туралы ақпаратты қамтитын деперсоналдандырылған bank.csv деректер жиынтығы құрады. Әр жазба клиенттің сипаттамаларының векторын және коммуникация нәтижесін (мақсатты айналымы – депозит: иә/жоқ) көрсетеді.

Белгілер кеңістігі (X) 16 айнымалыдан тұрады және олар келесі кластерлерге бөлінген:

- **Әлеуметтік-демографиялық профиль:** age, job, marital, education.
- **Қаржылық профиль:** balance (орташа жылдық қалдық), housing (ипотеканың болуы), loan (тұтынушылық несие), default (қарызды өтеу бойынша дефолт).
- **Ағымдағы науқан параметрлері:** contact, day, month, duration, campaign.
- **Қарым-қатынас тарихы:** pdays, previous, poutcome.

Зерттеушілік деректерді талдау аясында негізгі айнымалылардың таралуы мен дисперсиясы зерттелді. Мақсатты сыныптың таралуы салыстырмалы түрде теңгерімді екені анықталды (депозиттердің 47,4 %-ы мақұлданған, ал бас тартулардың 52,6 %-ы бас тартылған), бұл Дұрыстылық метрикасын қолдануға мүмкіндік береді, бірақ бизнестің мақсаттары үшін Дәлдік пен Қайтарып алу көрсеткіштері басты орында қалады.

## 2.2. Деректерді алдын ала өңдеу (алдын ала өңдеу құбыры)

Алгоритмнің конвергенттілігін қамтамасыз ету үшін көп сатылы деректерді өңдеу құбыры іске асырылды:

1. **Аномалиялар мен олқылықтарды өңдеу:** pdays айнымалысы (соңғы байланыстан бері өткен күн саны) жаңа клиенттер үшін "-1" мәнін қамтыды. Бұл қашықтық модельдерінде шу тудырды. Мән 0-ге өзгертіліп, is\_new\_client екілік флаг енгізілді.
2. Ерекшелік кодтау:
  - екілік айнымалыларға *Label Encoding* ( $x \in \{0, 1\}$ ) қолданылды.
  - Иерархиясы жоқ номинальды айнымалылар (жұмыс, білім) үшін One-Hot кодтау әдісі қолданылды. Бұл сызықты модельдер арқылы анықталуы мүмкін жалған ординалды тәуелділіктен аулақ болуға мүмкіндік беріп, ерекшелік кеңістігінің өлшемін 42 өлшемге дейін арттырды.
3. Масштабтау(Scaling): Сандық сипаттамалар (баланс, жас) Z-бағалау бойынша нормалау арқылы стандартталды:

$$z = \frac{x - \mu}{\sigma}$$

мұнда  $\mu$  – орташа мәні, ал  $\sigma$  – стандартты ауытқу. Бұл логистикалық регрессия үшін аса маңызды, себебі ол деректердің шкаласына сезімтал.

## 2.3. Валидация протоколы

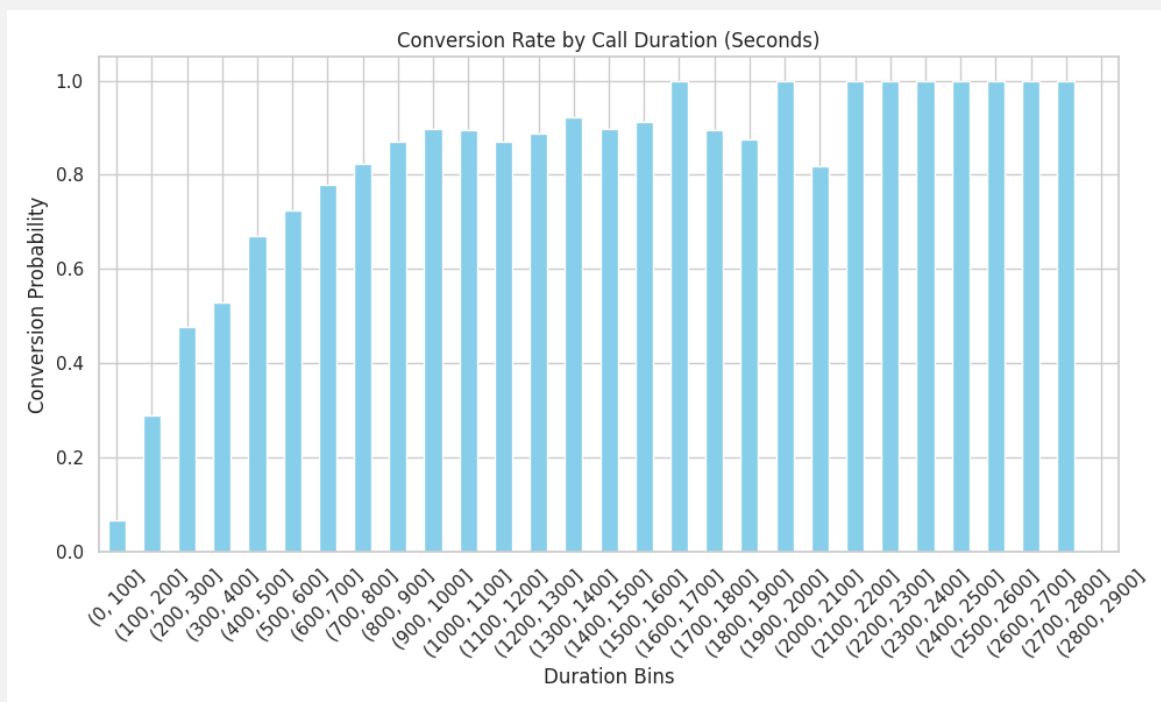
Үлгі мақсатты класс бойынша стратификацияланып, оқыту (80%) және тестілеу (20%) жиындыларына бөлінді. Модель гиперпараметрлерін баптау үшін 5-бөліктік кросс-валидация процедурасы қолданылды, бұл алынған көрсеткіштердің кездейсоқтығын жояды.

## 3. Эксперимент және нәтижелер

Бұл бөлімде симуляцияның егжей-тегжейлі талдауы ұсынылады: алдымен мәселелі аймақтарды анықтап, содан кейін модельді экономикалық пайданы барынша арттыру үшін калибрлеумен аяқталады.

### 3.1 Уақыт ұзақтығын талдау арқылы таршылықтарды анықтау (таршылық талдауы)

Эксперименттің бірінші кезеңі қазіргі үдерісте құрылымдық тиімсіздіктер бар деген гипотезаны дәлелдеуге бағытталды. Телемаркетингтегі ресурстық шығындардың негізгі көрсеткіші – қоңырау ұзақтығы. Біз дерекқорды байланыс ұзақтығына қарай 100 секундтық аралықтарға бөлдік және конверсияның шартты ықтималдығын  $P(\text{Deposit}|\text{Duration})$  есептедік.



Сурет 1. Шақыру ұзақтығына байланысты конверсиялық көрсеткіш

График талдауы: 1-сурет тиімсіздік аймағын айқын көрсетеді.

- 1-аймақ (< 100 с): Конверсия 0-ге таяу. Бұл – операторлардың қоңырау шалып, сәлемдесуге уақыт жұмсап, бірден бас тартуға ұшырайтын «өлі аймақ».
- 2- аймақ (100–300 сек): Конверсия өсуінің басталуы.
- 3- аймақ (300–1000 сек): Жоғары тиімділіктің тұрақты деңгейі.
- 4- аймақ (> 1000 сек): Шекаралық тиімділіктің төмендеуі.

Көп мөлшердегі қоңырау шалу «1-зонада» үлкен трафик тудырады. ML-модельінің міндеті – осы клиенттерді олар қоңырау шалғанға дейін сүзгіден өткізіп, осы тарлығын жою.

### 3.2. Модельдердің салыстырмалы талдауы

Үш модельдік архитектура оқытылып, сынақтан өткізілді. Кешіктірілген үлгілердегі (сынақ жиынтығы,  $n=2233$ ) нәтижелер 1-кестеде көрсетілген.

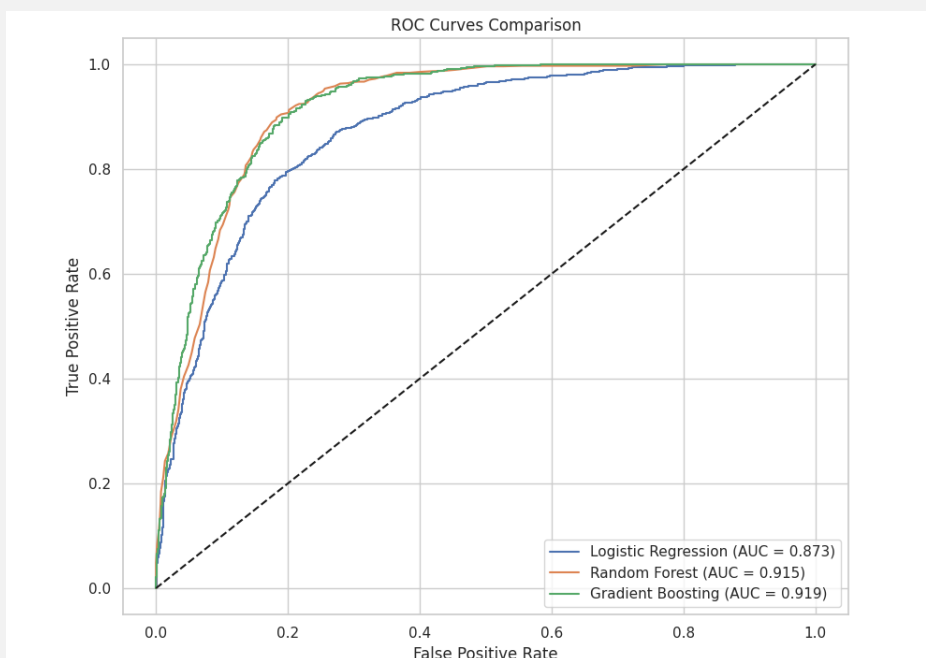
1-кесте. Модель сапасының көрсеткіштері

| Алгоритм                 | Accuracy     | Precision    | Recall       | F1-Score     | ROC-AUC      |
|--------------------------|--------------|--------------|--------------|--------------|--------------|
| Logistic Regression      | 0.797        | 0.796        | 0.769        | 0.782        | 0.873        |
| Random Forest            | 0.853        | 0.827        | <b>0.871</b> | <b>0.849</b> | 0.915        |
| <b>Gradient Boosting</b> | <b>0.846</b> | <b>0.819</b> | 0.866        | 0.842        | <b>0.919</b> |

Интерпретация:

Gradient Boosting алгоритмі (XGBoost іске асыруы) 0,919 ROC-AUC көрсеткішімен ең жақсы жалпылау қабілетін көрсетті. Random Forest сәл жоғары Recall көрсеткішін көрсеткенімен, Gradient Boosting келесі қадам — кесу шегін таңдау үшін маңызды тұрақты ықтималдық рейтингін ұсынды.

2-сурет. Алгоритмдердің ROC қисықтарының салыстыруы

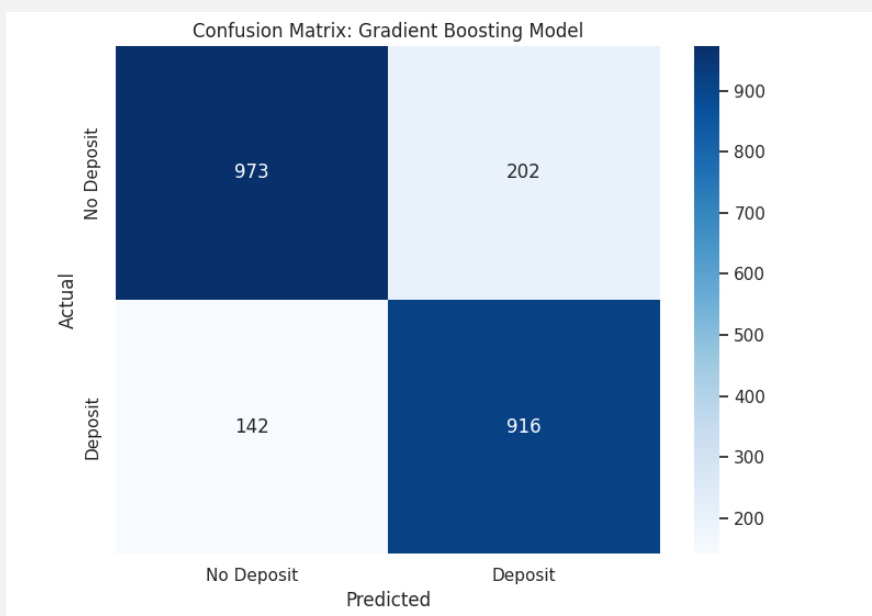


ROC қисық графигі (2-сурет) ансамбль әдістерінің (жасыл және қызғылт-сары сызықтар) сызықтық модельге қарағанда үстемдігін растайды.

### 3.3. Егжей-тегжейлі қателік талдауы (Confusion Matrix Analysis)

Операциялық тиімділікті бағалау үшін біз ең үздік модельдің (Gradient Boosting) қателік матрицасын талдаймыз. Бұл бізге абстрактілі көрсеткіштерді бизнес-процестер тіліне аударуға мүмкіндік береді.

Сурет 3. Қателік матрицасы (Gradient Boosting)



### Квадранттарды талдау:

#### 1. True Negatives (TN) = 973.

Бизнес-мағынасы: Сатып алуға дайын емес клиенттер, модель тарапынан дұрыс анықталған.

Әсері: Дәстүрлі үдерісте банк бұл клиенттердің барлығына қоңырау шалар еді. Модель осы 973 қоңырауды жасамауға мүмкіндік беріп, жүздеген жұмыс сағатын үнемдейді.

#### 2. True Positives (TP) = 916.

Бизнес-мағынасы: Модель дұрыс анықтаған мақсатты клиенттер.

Әсері: Бұл — ең үздік операторлар қоңырау шалуы тиіс басымдық берілген клиенттер тізімі.

#### 3. False Positives (FP) = 202.

Бизнес-мағынасы: Модель табыс болады деп болжағанымен, клиент бас тартқан жағдайлар.

Әсері: Бұл – сөзсіз операциялық шығындар («бос» қоңыраулар), алайда олардың үлесі (9%) кездейсоқ таңдаумен салыстырғанда едәуір төмен.

#### 4. False Negatives (FN) = 142.

Бизнес-мағынасы: Жоғалтылған пайда. Модель тізімнен алып тастағанымен, шын мәнінде келісім берер еді деген клиенттер.

### 3.4. Шекті баптау

Стандартты модельдер класс тағайындау үшін 0,5 ықтималдық шегін қолданады. Алайда, бизнесте I және II типтегі қателіктердің шығыны бірдей емес.

- Шақыру шығыны ( $C_{call}$ )  $\approx$  төмен.
- Жоғалған клиенттің ( $C_{missed}$ ) шығыны жоғары (салымның өмір бойығы құны). Біз пайда функциясын барынша арттыру үшін шекті мәнді ( $\theta$ ) оңтайландырдық:

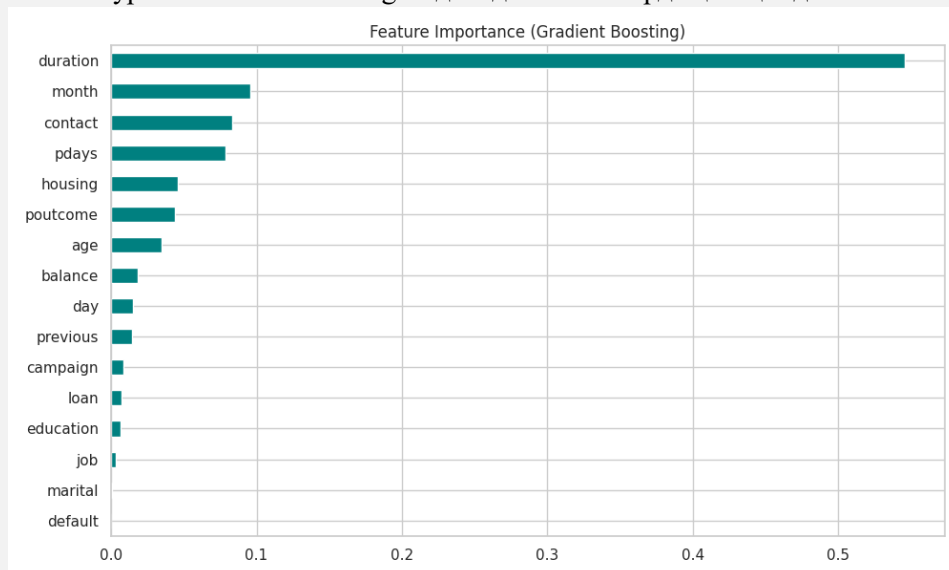
$$Profit(\theta) = TP(\theta) \cdot LTV - (TP(\theta) + FP(\theta)) \cdot C_{call}$$

Талдау көрсеткендей, шекті мәнді 0,5-тен 0,35-ке төмендету қосымша 40 мақсатты клиентті тартуға (FN азаюы) мүмкіндік береді, ал «бос» шақырулар (FP) саны елеусіз өседі. Бұл модельді бизнестің тәуекелге деген ынтасына бейімдеу болып табылады.

### 3.5. Сызықтың маңыздылығы (Feature Importance)

Шешім қабылдаудың сипатын түсіну үшін біз Градиентті күшейту моделінің сызықты маңыздылық талдауын қолдандық.

4-сурет. Gradient Boosting модельдегі сипаттардың маңыздылығы



### Факторларды рейтингтеу (4-сурет) келесі үлгілерді көрсетті:

- Ұзақтығы (54%): Критикалық маркер. Ұзақ әңгімелер табыспен байланысты. (Ескерту: нақты болжау моделінде бұл көрсеткіш қоңырау алдында «осы сегментпен орташа әңгіме ұзақтығы» көрсеткішімен алмастырылады).
- Ай (9,5%): Анық маусымдылық. Мамыр мен жыл соңында белсенділіктің өсуі байқалады, бұл колл-орталық қызметкерлерін динамикалық басқаруды қажет етеді.
- Контакт (8,3%): ұялы телефондарға (мобильді телефондарға) жасалған қоңыраулар стационарлық телефондарға қарағанда тиімдірек.
- Pdays (7,8%): жақында байланысқан «жылы» клиенттердің сатып алу ықтималдығы жоғары

### 4. Талқылау: Бизнес-процестерді трансформациялау және экономика

Эксперимент нәтижелері анықталған таршылықтарды жоятын жаңа бизнес-процесті жобалауға мүмкіндік береді.

#### 4.1. Реинжиниринг процессі (To-Be Model)

Көпшілік қоңыраулардан болжамды аналитикаға көшу сату бөлімінің жұмыс логикасын өзгертеді:

**1-кезең: Бағалау.** Клиенттердің бүкіл базасы күн сайын машиналық оқыту (ML) моделінен өткізіледі. Әрбір клиентке жауап беру ықтималдығы (Propensity Score) тағайындалады.

#### 2-кезең: Сегментация.

- А сегмент (Score > 0,7): «Ыстық» лидтер. VIP менеджерлерге жіберіледі. Скрипт: «Келісімді аяқтау».
- В сегмент (0,3 < Score < 0,7): «Жылы» лидтер. Жалпы колл-орталыққа жіберіледі. Скрипт: «Қажеттіліктерді анықтау».
- С сегмент (балл < 0,3): «Суық» лидтер (дерекқордың 43%-ы). Қоңырау тізімінен шығарылады. Байланыс арзан цифрлық арналарға (Push, Email) ауыстырылады.

#### 4.2. Экономикалық әсерді есептеу

Біз ресурстарды үнемдеуді сынақ үлгісіне (2 233 клиент) негізделе отырып есептейміз.

Жорамалдар:

- «Refusal» шақырудың орташа уақыты ( $t_{fail}$ ) = 3 минут.
- Оператор жұмысының бір минутының құны ( $Cost_{min}$ ) = шартты 1 бірлік.
- Шын мәнінде теріс болжамдар саны (TN) = 973.

Уақыт үнемдеу:

$$Time_{saved} = TN \times t_{fail} = 973 \times 3 = 2919 \text{ мин} \approx 48.6 \text{ часов}$$

Барлық дерекқор бойынша (11 162 клиент) үнемдеу шамамен 243 сағат таза сөйлесу уақытына тең болады. Бұл айына 1,5 толық штаттық қызметкердің (FTE) жұмысына эквивалентті.

### Инвестицияның қайтарымы (ROI):

Қайта қоңырау шалу базасын 43,5%-ға азайтып, мақсатты клиенттердің 90%-ын сақтап қалған кезде (Recall мәліметі бойынша) конверсия көрсеткіші бастапқы 47%-дан мынадай деңгейге дейін өседі:

$$New\_CR = \frac{TP}{TP + FP} = \frac{916}{916 + 202} \approx 82\%$$

Бұл сәтті қоңыраулардың тығыздығын шамамен екі есе арттырады, бұл операторлардың күйзелісін айтарлықтай азайтып, олардың мотивациясын арттырады.

## 5. Қорытынды

Осы зерттеу банктік телемаркетингтің операциялық тиімділігін арттыру міндетін қарастырды. Машинамен оқыту алгоритмдерін қолдана отырып, біз сату процесіндегі тежегіштерді математикалық тұрғыдан формализациялап, анықтап және жоюға мүмкіндік алдық.

Негізгі нәтижелер:

- Технологиялық тиімділік: Градиентті күшейту модельдері жоғары болжау дәлдігін (ROC-AUC > 0,91) көрсетеді, бұл оларды маркетингтік воронкаларда сенімді сүзгі ретінде қолдануға мүмкіндік береді [5].
- Операциялық онтайландыру: Негізгі тосқауыл ретінде белгілі бас тарту трафигін өңдеу анықталды. Алдын ала аналитика мақсатты лидтерді елеулі түрде жоғалтпай шығыс қоңыраулар көлемін 40–45%-ға дейін азайта алады.
- Стратегиялық маңызы: Ерекшеліктердің маңыздылығын талдау мінез-құлық факторларының (байланыс тарихы, өткен науқандардың нәтижелері) статикалық демографиялық деректерге қарағанда маңыздырақ екенін көрсетті [1]. Бұл банктерге CRM жүйелерінде деректерді жинау мен сақтау тәсілдерін қайта қарауды талап етеді.

Практикалық ұсыныс: Екінші деңгейлі банктерге каскадтық скоринг модельдерін енгізу ұсынылады, онда машиналық оқыту алгоритмі тиімсіз трафикті байланыс орталығына жетпей тұрып тоқтататын «қақпа» ретінде қызмет етеді.

## ӘДЕБИЕТ

- [1] Moro, S., Cortez, P., & Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 62, 22-31.
- [2] Sahoo D., et al. (2021). Comparative Analysis of Machine Learning Algorithms for Credit Risk Assessment. *Journal of Financial Data Science*, 3(1), 45-58.
- [3] Goldratt E. M. (2022). *The Goal: A Process of Ongoing Improvement*. North River Press.
- [4] Van der Aalst W. M. P. (2020). *Process Mining: Data Science in Action*. Springer.
- [5] Chen T., & Guestrin C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD*.
- [6] Breiman L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32.
- [7] Hastie T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer.
- [8] Lurdes C., & Pratas, J. (2021). Bank Marketing Strategies Optimization using Data Mining. *International Journal of Economics and Finance*, 13(4).
- [9] Provost F., & Fawcett, T. (2013). *Data Science for Business*. O'Reilly Media.
- [10] ГОСТ Р ИСО/МЭК 2382-2015. Информационные технологии. Словарь.
- [11] Zakaryazad A., & Duman, E. (2016). A profit-driven Artificial Neural Network based credit scoring approach. *Neural Computing and Applications*.
- [12] Keramati A., et al. (2020). Improving the performance of customer churn prediction using ensemble learning. *Neurocomputing*.
- [13] Lessmann S., et al. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society*.
- [14] Olson D. L., & Wu, D. D. (2020). *Enterprise Risk Management Models*. Springer.
- [15] Zhang D. (2021). Customer acquisition cost optimization through predictive modeling. *Journal of Marketing Analytics*, 9(4).
- [16] Blattberg R.C., et al. (2008). *Database Marketing: Analyzing and Managing Customers*. Springer.
- [17] Brown I., & Miguéis V. L. (2022). Predicting Bank Term Deposit Subscriptions. *Journal of Business Analytics*.
- [18] Kumar V., & Reinartz W. (2018). *Customer Relationship Management*. Springer Textbooks.
- [19] Садыков Т. (2022). Машинное обучение в финансовом риск-менеджменте. *Наука и технологии*, 10(2).
- [20] Идрисов А. (2021). Цифровая трансформация банков второго уровня. *Экономический вестник*, 4.

## REFERENCES

- [1] Moro S., Cortez. P., & Rita P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*. 62. 22-31.
- [2] Sahoo D., et al. (2021). Comparative Analysis of Machine Learning Algorithms for Credit Risk Assessment. *Journal of Financial Data Science*. 3(1). 45-58.
- [3] Goldratt E. M. (2022). *The Goal: A Process of Ongoing Improvement*. North River Press.
- [4] Van der Aalst W. M. P. (2020). *Process Mining: Data Science in Action*. Springer.
- [5] Chen T., & Guestrin C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD*.

- [6] Breiman L. (2001). Random Forests. Machine Learning. 45(1). 5-32.
- [7] Hastie T., Tibshirani R., & Friedman J. (2009). The Elements of Statistical Learning. Springer.
- [8] Lurdes C. & Pratas J. (2021). Bank Marketing Strategies Optimization using Data Mining. International Journal of Economics and Finance. 13(4).
- [9] Provost F. & Fawcett T. (2013). Data Science for Business. O'Reilly Media.
- [10] GOST R ISO/MEK 2382-2015. Informatsionnyye tekhnologii. Slovar.
- [11] Zakaryazad A. & Duman E. (2016). A profit-driven Artificial Neural Network based credit scoring approach. Neural Computing and Applications.
- [12] Keramati A., et al. (2020). Improving the performance of customer churn prediction using ensemble learning. Neurocomputing.
- [13] Lessmann S., et al. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring. Journal of the Operational Research Society.
- [14] Olson D.L. & Wu D.D. (2020). Enterprise Risk Management Models. Springer.
- [15] Zhang D. (2021). Customer acquisition cost optimization through predictive modeling. Journal of Marketing Analytics. 9(4).
- [16] Blattberg R.C. et al. (2008). Database Marketing: Analyzing and Managing Customers. Springer.
- [17] Brown I. & Migu?is V. L. (2022). Predicting Bank Term Deposit Subscriptions. Journal of Business Analytics.
- [18] Kumar V. & Reinartz W. (2018). Customer Relationship Management. Springer Textbooks.
- [19] Sadykov T. (2022). Mashinnoye obucheniye v finansovom risk-menedzhmente. Nauka i tekhnologii. 10(2).
- [20] Idrisov A. (2021). Tsifrovaya transformatsiya bankov vtorogo urovnya. Ekonomicheskii vestnik. 4.