

УНИВЕРЗИТЕТ У БЕОГРАДУ  
МАТЕМАТИЧКИ ФАКУЛТЕТ



Александра Б. Поповић

СТАТИСТИЧКО МОДЕЛОВАЊЕ  
УЗРОЧНО-ПОСЛЕДИЧНИХ ВЕЗА СА  
ПРИМЕНАМА У БАНКАРСТВУ

мастер рад

Београд, 2023.

**Ментор:**

проф. др Бојана МИЛОШЕВИЋ, ванредни професор  
Универзитет у Београду, Математички факултет

**Чланови комисије:**

др Марија ЦУПARIЋ, доцент  
Универзитет у Београду, Математички факултет

др Марко ОБРАДОВИЋ, доцент  
Универзитет у Београду, Математички факултет

**Датум одбране:** септембар 2023.

Захваљујем се професорки Бојани Милошевић на помоћи, подршци и стрпљењу током писања овог рада.

Захвалила бих се и осталим члановима комисије, професору Марку Обрадовићу и професорки Марији Цупарић, на стрпљењу приликом читања рада и свим корисним сугестијама.

Захвалност дугујем и својим родитељима који су увек били несебична подршка.

Посебну захвалност дугујем свом најбољем другу Михајлу за подршку током читавих студија и безброј заједно спремљених испита.

# Садржај

<b>1</b>	<b>Увод</b>	<b>1</b>
1.1	Моделовање уздизања . . . . .	2
1.2	Примена у банкарству . . . . .	3
<b>2</b>	<b>Основе узрочно-последичног закључивања</b>	<b>5</b>
2.1	Рандомизирани контролни тестови . . . . .	8
2.2	Пример рачунања вредности ефекта третмана . . . . .	10
<b>3</b>	<b>Моделовање уздизања</b>	<b>14</b>
3.1	Дефинисање циљне променљиве . . . . .	15
3.2	Приступ два модела . . . . .	16
3.3	Приступ трансформације класа . . . . .	17
3.4	Приступ директног моделовања уздизања . . . . .	21
<b>4</b>	<b>Евалуација</b>	<b>24</b>
4.1	Традиционалне статистике уздизања . . . . .	25
4.2	Метрике базиране на $Y^*$ . . . . .	30
<b>5</b>	<b>Преглед библиотеке Causal ML и Meta-Learner алгоритама</b>	<b>32</b>
5.1	Meta-Learner алгоритми . . . . .	33
<b>6</b>	<b>Примена модела уздизања на реалне податке</b>	<b>38</b>
6.1	Упознавање са подацима . . . . .	40
6.2	Припрема података . . . . .	41
6.3	Модел склоности . . . . .	41
6.4	Модели . . . . .	42
6.5	Предвиђање модела . . . . .	46
<b>7</b>	<b>Закључак</b>	<b>55</b>

*САДРЖАЈ*

v

**Библиографија**

**57**

# Глава 1

## Увод

Узрочно-последичне везе као појам се односе на проучавање и разумевање узрока и последица између различитих променљивих у системима. Област која се бави моделирањем узрочно-последичних веза се фокусира на идентификовање каузалних ефеката, односно утврђивање како једна променљива утиче на другу.

Узрочно-последичне везе су кључне за разумевање многих феномена у пословном свету, медицини, социологији, па тако и у банкарству. У прошлости, статистички модели су се углавном фокусирали на предвиђање коначних исхода на основу варијабли које су описивале опште карактеристике популације. Међутим, у многим ситуацијама није довољно само предвидети крајњи исход, потребно је разумети узрочно-последичне везе између одређених варијабли и исхода.

У том смислу, недавни развој статистичких модела каузалности (тј. узрочно-последичних веза) омогућио је разумевање и моделирање узрочности на начин који раније није био могућ. Пример таквог модела, о ком ће бити речи у овом раду, се дефинише у приступу званом *Моделовање уздизања* (енг. Uplift modeling).

Модели уздизања се фокусирају на израчунавање ефеката третмана, тј. очекиване разлике у коначном исходу између третираних и нетретираних група. Ово је корисно у ситуацијама када третман има позитиван ефекат само на одређене групе, док на неке друге групе може имати негативан ефекат, или га уопште нема.

Овај рад ће представити преглед литературе о моделовању уздизања, и размотриће три различите технике за оцену узрочних ефеката: метод два

модела, приступ трансформације класе и модификацију постојећих модела машинског учења. Евалуација квалитета модела биће представљена у 4. поглављу, при чему ће бити приказане децилна анализа и криве уздизања, а такође, биће описане и мере квалитета попут средњеквадратне грешке.

Поред тога, биће речи о примени моделовања уздизања у пословном контексту, примарно у банкарству, на конкретним подацима где ћемо поред поменутих техника увести још три алгорита за оцену узрочних ефеката, тако зване мета алгоритме.

## 1.1 Моделовање уздизања

Статистички модели узрочно-последичних веза представљају скуп техника које се односе на испитивање допунског утицаја неке акције или третмана на крајњи исход корисника, нпр. скуп техника које нека компанија или појединац може да користи за процену повећања потрошње клијената тј. испитује се ефекат неке акције на исход корисника.

Ове информације могу бити корисне компанијама приликом одабира клијената којима ће бити понуђен неки производ или услуга. На пример, један од интереса банке је повећање броја клијената и зато је важно пронаћи скуп мера које могу допринети томе. Како је потребно проценити разлику између два исхода који се међусобно искључују за појединца (особа добија неки третман, нпр. добија понуду од компаније или је не добија), ослањамо се на рандомизирани експерименте, тј. насумично додељивање купаца тј. клијената да добију третман (третман група) или не добију третман (контролна група). Тиме овај проблем постаје проблем моделовања узрочно-последичних ефеката. Последњих година се за овај проблем користе разне методе машинског учења.

Да би се то остварило, потребно је обучити различите моделе и изабрати онај који даје најпоузданије предвиђање према неким метрикама учинка.

Приступ моделовања уздизања је подељен у три главна приступа:

- Приступ два модела (енг. Two-Model approach);
- Приступ трансформације класа (енг. Class Transformation approach);
- Приступ директног моделовања (енг. Directly uplift modeling).

Користићемо нотацију модерне економетрије, по узору на модел из рада Рубинова (1974.), како бисмо пружили јасније поређење и евалуацију ова три приступа и генерализовали један од њих.

У наставку рада ћемо показати да је минимизирање средњеквадратне грешке ( $MSE$ ) у односу на процену узрочно-последичног ефекта, еквивалентно минимизацији модификоване средњеквадратне грешке ( $MSE^*$ ).

Поред поменутих три приступа за оцену ефекта третмана, истим проблемом се баве и мета модели који користе различите моделе и стратегије процене ових ефеката. Мета-Модели су општа класа алгоритама који се користе за узрочно закључивање и моделовање ефекта уздизања, а поменутих три приступа (приступ са два модела, приступ трансформације класа и модификација постојећих модела) су специфичне технике унутар овог оквира које пружају различите стратегије за процену ефеката третмана.

Пре него што то покажемо, осврнућемо се на главне примене области узрочно-последичних веза у банкарству.

## 1.2 Примена у банкарству

Статистички модели узрочно-последичних веза се могу применити у банкарству на различите начине. Један од главних циљева примене ових модела у банкарству је идентификација узрочних фактора који доприносе побољшању пословања банке и повећању њених прихода. На пример, узрочно-последични модели могу се користити за идентификацију клијената који су склонији да се одазову на промотивне понуде банке или да користе одређене производе и услуге банке. Знајући такве информације, послодавци би били у могућности да ефикасно класификују и циљају купце.

Ови модели такође могу бити коришћени за изградњу персонализованих препорука производа и услуга које су прилагођене индивидуалним потребама клијената. Осим тога, узрочно-последични модели могу се користити за анализу ризика и процену кредитне способности клијената, као и за предвиђање потенцијалних губитака и штета у случају неплаћања кредита или других финансијских обавеза.

Узрочно-последични модели такође могу бити корисни у анализи ефикасности маркетиншких кампања банке, што може довести до бољег разумевања утицаја маркетиншких активности на повећање прихода и профитабилности



банке. Укратко, примена узрочно-последичних модела у банкарству може помоћи банкама да побољшају своје пословање, повећају приходе и профитабилност, и боље разумеју своје клијенте и тржиште.

## Глава 2

# Основе узрочно-последичног закључивања

Рубинов модел узрочности, објављен 1974. године, представља један од најчешће коришћених модела за разумевање узрочних ефеката. Такође познат као *потенцијални модел резултата*, представља статистички модел који се користи за процену узрочних ефеката у оквиру експерименталних и неексперименталних истраживања. Основна идеја модела су *потенцијални резултати*, који представљају тумачење, шта би се догодило догодило са појединцима у ситуацијама третмана или алтернативним сценаријима.

Циљ модела је да процени ефекат третмана на индивидуалном нивоу, који се може разликовати у зависности од карактеристика појединца. Ово је посебно корисно када су у питању неекспериментална истраживања, где није могуће рандомизирати третмане на начин који би обезбедио еквивалентне групе, јер се могу узети у обзир разлике између групе које се посматрају, а које нису последица третмана, и тако проценити стварни узрочни ефекат третмана.

У основи Рубиновог модела су појмови потенцијалних исхода и узрочних ефеката. Посматрамо експеримент са  $N$  појединаца индексираних са  $i$ . Означавамо са  $Y_i(1)$  исход појединца  $i$  када прими одређени третман, а са  $Y_i(0)$  исход појединца  $i$  када не прима третман. Тада дефинишемо узрочни ефекат  $\tau_i$ , активног третмана у поређењу са контролним третманом као:

$$\tau_i = Y_i(1) - Y_i(0). \quad (2.1)$$

У истраживањима се често помиње вредност *CATE* (енг. Conditional

Average Treatment Effect), тј. условни просечни ефекат третмана и његова процена. Ова величина представља очекивани узрочни ефекат третмана на подскуп целе популације:

$$CATE : \hat{\tau}(X_i) = E[Y_i(1)|X_i] - E[Y_i(0)|X_i]. \quad (2.2)$$

Величина  $X_i$  представља вектор случајних варијабли, димензије  $L \times 1$ .

Није могуће истовремено посматрати обе случајне променљиве  $Y_i(1)$  и  $Y_i(0)$ . Зато дефинишемо  $W_i \in \{0, 1\}$  као бинарну варијаблу која има вредност 1 ако особа  $i$  прима активни третман, и има вредност 0 ако особа  $i$  прима контролни третман. Посматрани исход особе  $i$  се дефинише као:

$$Y_i^{obs} = W_i Y_i(1) + (1 - W_i) Y_i(0). \quad (2.3)$$

Популарно али нажалост нетачно уверење је да се *CATE* увек може оценити из посматраних (опсервацијских) података, једноставним израчунавањем емпиријског еквивалента (статистичка мера која се добија из емпиријских података и која је аналогна теоријској мери) следећег израза:

$$E[Y_i^{obs}|X_i = x, W_i = 1] - E[Y_i^{obs}|X_i = x, W_i = 0]. \quad (2.4)$$

Овај израз неће оценити *CATE* из опсервацијских података само применом формуле, осим ако не уведемо следеће претпоставке:

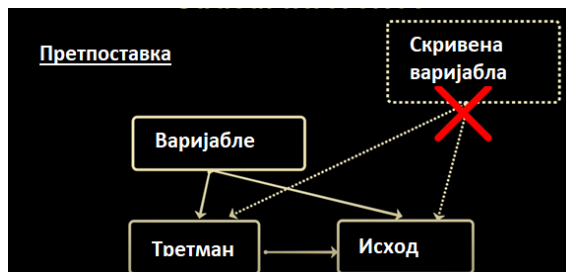
### 1. Претпоставка о неосујећености<sup>1</sup>

У процесу узрочно-последичног закључивања, ова претпоставка се односи на то да не постоји скривени или неконтролисани фактор тј. варијабла која би могла искривити везу између третмана и исхода (Слика 2.1). Другим речима, желимо осигурати да је третман (интервенција) насумично додељен и некорелиран са свим другим факторима који утичу на исход. Ако је ова претпоставка испуњена, то значи да су третирана и нетретирана група еквивалентне по свим факторима, осим по самом третману, што омогућава тачно мерење ефекта третмана. Ова претпоставка је задовољена када се додељивање третмана врши насумично и условљено је са  $X_i$  тј. узимамо у обзир карактеристике појединаца, и њена претпоставка је кључна за валидну процену каузалних

---

<sup>1</sup>енг. Unconfoundedness Assumption

ефеката третмана из посматраних података. Ова претпоставка се често постиже применом метода рандомизације како би се минимизирао утицај збуњујућих фактора.



Слика 2.1: Претпоставка о неосујећености

## 2. Претпоставка о условној независности<sup>2</sup>

Ова претпоставка се односи на то да је третман независан од последице ако се узму у обзир све релевантне варијабле које могу утицати на њихов однос. Ово значи да не постоји разлика у карактеристикама између третиране и нетретиране групе које би могле објаснити разлику у исходу. Ако је ова претпоставка испуњена, то значи да нема збуњујућих фактора (енг. confounding factors) који би искривили процену ефекта третмана.

$$CIA : Y_i(1), Y_i(0) \perp W_i | X_i$$

## 3. Независност третмана<sup>3</sup>

То је претпоставка која се односи на независност третмана који се примењује на једну јединицу (нпр. појединца) од третмана који се примењује на друге јединице. Другим речима, претпоставља се да третман који је додељен једној јединици не утиче на третман или исход других јединица. Ова претпоставка је важна јер омогућава изоловање ефеката самог третмана на исход, без утицаја интеракција између третираних јединица.

## 4. Пристрасност<sup>4</sup>

Битно је да третманска група буде добра репрезентација целе попула-

<sup>2</sup>енг. Conditional Independence Assumption (CIA)

<sup>3</sup>енг. SUTVA - Stable unit treatment value assumption

<sup>4</sup>енг. Selection bias

ције, тј. да су клијенти униформно распоређени у обе групе по својим карактеристикама.

Важно је напоменути да ове претпоставке могу бити изазовне за остваривање у пракси, поготово у сложеним системима. Стога процена и потврда ових претпоставки захтевају пажљиво разматрање и примену адекватних статистичких метода.

## 2.1 Рандомизирани контролни тестови

Најчешћи случајеви на којима примењујемо моделе уздицања јесу управо рандомизирани контролни тестови, у пракси познатији као А/Б тестови. Циљеви експеримента могу бити различити, у зависности од врсте проблема. За илустрацију примера у банкарству, нека у нашем случају циљ експеримента буде повећање продаје у банци.

Тест се зове рандомизирани из разлога што се клијенти случајним избором распоређују у третирану и контролну групу. На овај начин смо сигурни да уколико донесемо закључак да је одабрани третман повећао продају код клијената, да је то искључиво због утицаја третмана а не због неких појединачних карактеристика клијената који су изабрани да буду у овим групама.

А/Б тестирање је експериментална метода која се користи у контексту истраживања и оптимизације у разним областима, и примењује се како би се упоредиле две (или више) варијације неког елемента или процеса, да би се одредило која даје боље резултате. Једна група клијената се излаже оригиналној верзији (контролна група, без третмана), док се друга група излаже третману. Обе групе се прате и анализирају како би се проценило како свака варијација утиче на резултате и перформансе.

Кључна предност А/Б тестирања је могућност да се добију емпиријски подаци о томе која варијација доноси боље резултате, уместо да се ослањамо само на претпоставке или субјективне процене. Поступак А/Б тестирања укључује следеће кораке:

1. **Идентификација циља:** Дефинише се јасан циљ експеримента, нпр повећање продаје.
2. **Дефинисање варијабли:** Идентификују се специфичне варијабле које ће бити тестиране, везане за клијенте.

3. **Подела корисника:** Корисници се насумично распоређују у контролну и третирану групу како би се изложили различитим третманима.
4. **Извођење експеримента:** Контролна група остаје изложена оригиналној верзији, док се третирана група излаже дефинисаном третману.
5. **Прикупљање података:** Прикупе се релевантни подаци о понашању корисника који нам дају информацију о акцији коју је корисник потенцијално преузео (нпр. у вези куповине).
6. **Анализа резултата:** Користе се статистички алати како би се анализирали резултати и утврдило да ли постоји значајна разлика између контролне и третиране групе.
7. **Закључци и имплементација:** На основу анализе резултата доносе се закључци о томе која варијација је ефикаснија и могу се имплементирати промене на основу тих закључака.

Важно је да А/Б тестирање буде добро осмишљено и изведено како би се осигурала валидност и поузданост резултата. То укључује довољан број учесника, праћење релевантних метрика и примену статистичких анализа.

## Пример А/Б тестирања

У банкарском сектору, А/Б тестирање се може применити како би се проценио утицај одређених промена на понашање корисника или клијената. Претпоставимо да банка жели да побољша ангажовање корисника на свом веб сајту и да повећа конверзије приликом отварања нових рачуна. Банка примењује А/Б тестирање како би проверила који дизајн веб странице доприноси бољим резултатима.

1. **Подела корисника:** Прво, банка насумично дели своје посетиоце веб сајта у две групе: контролну групу (А група) и третирану групу (Б група). Обе групе ће бити изложене веб сајту са различитим дизајном, при чему ће се једна група суочити са старим дизајном, а друга са новим дизајном.
2. **Примена промене:** Банка сада примењује промену на веб сајту за третирану групу (Б групу). Ово може бити било каква промена, као што је

другачији распоред елемената на страници, другачије боје, побољшана навигација или додатак позива на акцију за отварање рачуна.

3. Праћење резултата: Банка прати понашање корисника у обе групе током одређеног периода. На пример, прати колико корисника је посетило страницу за отварање рачуна, колико је заиста отворило рачун, колико је корисника напустило страницу без отварања рачуна, итд.
4. Анализа резултата: Након што је довољно података прикупљено, банка анализира резултате. Упоредиће перформансе обе групе и утврђује да ли је промена дизајна веб сајта (третман) довела до повећања конверзија приликом отварања нових рачуна у односу на стару верзију (контрола).

На основу резултата А/Б тестирања, банка може донети информисане одлуке о томе да ли ће задржати нови дизајн са побољшањима или ће се вратити на стари. Овај приступ А/Б тестирању омогућава банци да сазнаје шта стварно функционише у побољшању корисничког ангажовања и конверзија, на основу стварних података и статистичких анализа. А/Б тестирање је метода за упоређивање две или више варијанти третмана, док је моделовање уздизања техника за процену индивидуалног ефекта третмана како би се идентификовали они који су највероватније позитивно реаговали на третман.

## 2.2 Пример рачунања вредности ефекта третмана

Након рандомизираних тестова, за разумевање модела уздизања поменућемо и приказати на практичном примеру следеће битне појмове: *појединачни ефекат третмана*, *просечни ефекат третмана* и *условни просечни ефекат третмана*.

Претпоставимо да имамо скуп јединки од којих су неке у третманској, а неке у контролној групи. Циљна променљива је бинарна варијабла са вредностима  $\{0, 1\}$  и представља неку дефинисану акцију коју меримо [9]. У следећој табели имамо пример од 10 клијената од којих су неки у контролној а неки у третманској групи.

Клијент	Године	Исход третмана	Исход контролне групе
1	26	1	
2	24		1
3	48	0	
4	35		1
5	25		0
6	39		0
7	51	1	
8	24	0	
9	67	1	
10	34		0

**Просечни ефекат третмана** (енг. average treatment effect, ATE) је статистичка мера која се користи у контексту каузалног закључивања и евалуације ефеката третмана. Дефинише се као просечна разлика у резултатима између групе која је изложена третману и групе која није изложена третману. Прецизније, ATE представља просечну разлику између очекиваних резултата (нпр. исхода, перформанси) за све јединице популације које су биле третиране и оне које нису биле третиране. Ова мера пружа информацију о укупном утицају третмана на популацију и омогућава процену ефикасности третмана. На нашем примеру, то се рачуна на следећи начин:

$$Mean(treatment) = \frac{1 + 0 + 1 + 0 + 1}{5} = +0.6$$

$$Mean(control) = \frac{1 + 1 + 0 + 0 + 0}{5} = +0.4$$

$$ATE = 0.6 - 0.4 = +0.2.$$

Посматрајући само вредност просечног ефекта третмана, рекли бисмо да третман има позитиван утицај на циљану популацију.

Међутим уколико у калкулацију урачунамо и варијаблу година, закључујемо да су просечне године клијената у третираној групи 48 година, а у контролној групи 29,5 година. Ова варијабла представља неконтролисану варијаблу која може пружити неки додатни утицај на крајњи резултат, па је потребна детаљнија анализа. За потребе примера, коришћењем алгоритама



машинског учења или других техника, импутоваћемо недостајуће вредности у табели<sup>5</sup>.

Једна од могућих техника импутације се назива упаривање (енг. matching). То се постиже идентификацијом друге јединице ( $x_1$ ) која је веома слична (према некој метрици) нашој одабраној јединици ( $x_0$ ), али припада другој третманској групи. У случају датог примера, сматрамо да су особе сличних година заправо сличне јединице. Проналажењем најближих суседа свакој јединици, можемо импутовати недостајуће исходе у случају супротног третмана који има његов сусед. На тај начин предвиђамо вредности које се нису десиле. Сада табела изгледа овако:

Клијент	Године	Исход третмана	Исход контролне групе	ITE
1	26	1	1	0
2	24	0	1	-1
3	48	0	0	0
4	35	1	1	0
5	25	0	0	0
6	39	1	0	1
7	51	1	0	1
8	24	0	1	-1
9	67	1	1	0
10	34	1	0	1

**Појединачни ефекат третмана** (енг. individual treatment effect, ITE) се рачуна за сваку особу појединачно као разлика третираног и контролног исхода. На основу тога се вредност ATE може израчунати као просечна вредност од ITE, те закључујемо да третман заиста има позитиван ефекат на овај експеримент.

$$ATE = \frac{-1 + 1 + 1 - 1 + 1}{10} = +0.1$$

**Условни просечни ефекат третмана** је статистичка мера која представља просечан ефекат третмана на зависну променљиву, уз услов да су

<sup>5</sup>Недостајућих вредности формално нема зато што свака особа може припадати тачно једној групи, али када посматрамо ефекат варијабле године на израчунавање ATE, уравни- тежићемо обе групе импутацијом недостајућих вредности.

познате или контролисане друге варијабле. Израчунава се као разлика у очекиваном резултату између групе која је изложена третману и групе која није, при чему се узимају у обзир специфични услови или карактеристике које су релевантне за анализу. Ова мера омогућава процену ефеката третмана у специфичним подгрупама и пружа дубље разумевање како третман утиче на појединачне јединице у различитим контекстима. У нашем примеру тај услов биће варијабла година. За потребе примера, нека је праг који посматрамо 35 година:

$$CATE(age \geq 35) = \frac{0 + 0 + 1 + 1 + 0}{5} = +0.4$$

$$CATE(age < 35) = \frac{0 - 1 + 0 + 0 - 1 + 1}{5} = -0.2$$

На основу израчунатих условних просечних ефеката третмана за ова два услова, закључујемо да је ефекат третмана позитиван код особа које имају више од 35 година, док код осталих особа нема потребе примењивати овај третман.

Разумевањем ових величина закључујемо да је за оцену ефекта третмана потребно укључити све релевантне факторе који описују дату јединку. У складу са овим сазнањима бисмо на паметан начин распоређивали јединке у контролну и третирану групу у будућим анализама.

## Глава 3

# Моделовање уздизања

Многе компаније у разним областима пословања су заинтересоване за моделовање уздизања како би процениле ефекат неке акције тј. третмана на исход корисника. У терминима наведене нотације, моделовање уздизања се своди на оцену *CATE* вредности.

Компаније могу лако да спроведу рандомизирани експерименти јер често имају велике базе корисника или клијената, што им омогућава да имају довољно испитаника за рандомизацију у различите групе. Већи број испитаника повећава снагу и поузданост експеримента, чиме се добијају тачнији резултати. Многе компаније имају онлајн платформе и алате који омогућавају laku имплементацију експеримента. Путем платформи и софтвера, могуће је насумично доделити кориснике у различите третмане и пратити њихово понашање. Иако компаније могу лако да осигурају да је претпоставка о условној независности (CIA) задовољена, чињеница да никада нећемо посматрати стварну вредност  $\tau_i$ , чини отежано коришћење стандардних алгоритама за надгледано учење за његову оцену. Ако бисмо претпоставили да је  $\tau_i$  заиста могуће израчунати, једноставно бисмо поделили податке на скуп за тренирање и тестирање и користили бисмо један од многих доступних алгоритама машинског учења да бисмо добили оцену *CATE*  $\hat{\tau}(X_i)$  која минимизира функцију губитка на тренинг подацима.

Под моделовањем уздизања, постоје 3 главна приступа за оцењивање  $\tau(X_i)$ . Први поменути приступ је приступ 2 модела који подразумева изградњу два предиктивна модела, где се један користи за групу клијената који су добили одређени третман, а други модел се користи искључиво за клијенте који нису добили третман. Други приступ је класна трансформација варија-

бле у случају бинарне варијабле, која је први пут представљена у раду [16]. Трећи приступ је директно моделовање уздицања кроз модификацију добро познатих алгоритама класификације машинског учења, као што су стабла одлучивања, случајне шуме или методе потпорних вектора.

### 3.1 Дефинисање циљне променљиве

Пре него што уведемо детаљније три поменута приступа, описаћемо како се дефинише циљна променљива.

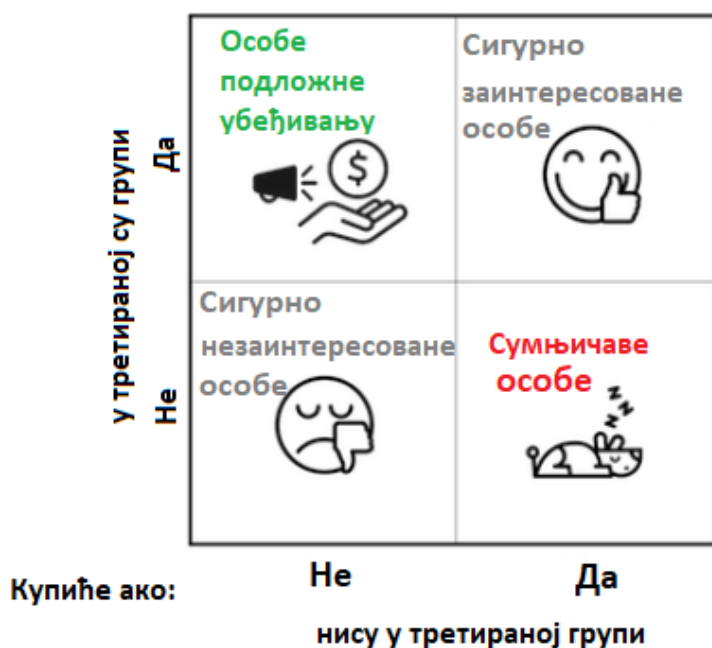
Као што је поменуто у претходном поглављу, циљ неког експеримента је да на основу одабраног третмана тј. акције коју предузимамо на скуп јединки у банци, повећамо одзив или поспешимо продају, или нешто треће у зависности од природе проблема. Компаније лако могу да спроведу рандомизирани контролни тестови (А/Б тестови) и на крају донесу финални закључак да ли је третман позитивно утицао на исход. Тај одговор важи на нивоу целог скупа података, али ми не знамо на индивидуалном нивоу да ли је третман успешан.

Клијенте распоређујемо у 4 одвојене групе:

1. **Особе подложне убеђивању (енг. Persuadables):** Ове особе ће купити производ уколико приме третман. Ова група људи је она коју желимо да таргетирамо. Уколико не приме третман, неће купити производ.
2. **Сигурно незаинтересоване особе (енг. Lost causes):** Особе које сигурно неће купити производ, независно од тога да ли се нађу или не нађу у третираној групи. На ове особе нема смисла трошити ресурсе.
3. **Сигурно заинтересоване особе (енг. Sure things):** Особе које ће сигурно купити производ, независно од тога да ли ће бити или неће бити у третираној групи. На ове особе нема смисла трошити ресурсе.
4. **Сумњичаве особе (енг. Sleeping dogs):** Особе које неће купити производ уколико се нађу у третираној групи, у супротном случају би купиле производ. Таргетирање ових особа је лоше по бизнис.

Поменута категоризација је уобичајен начин за класификацију појединаца у моделирању уздицања на основу њиховог потенцијалног одговора на трет-

ман (Слика 3.1<sup>1</sup>). Ове категорије су корисне за разне компаније у доношењу одлука за разне маркетиншке стратегије, где ће се фокусирати на циљање само одређених појединаца.



Слика 3.1: Категоризација клијената у 4 групе

Према подели је јасно да је нама у интересу да таргетирамо само групу особа **подложних убеђивању**. Таргетирање конкретне групе људи тј. појединаца је ефикасније и чува новац и ресурсе. Задатак модела је да за почетак одреди ову групу људи.

## 3.2 Приступ два модела

Приступ два модела је први од три приступа моделовања уздицања и најчешће се користи као основни модел. Овај приступ је представљен у следећим радовима [21] и [22], а у последње време се користи и за новије моделе машинског учења као што је алгоритам два стабла (енг. Two-Tree) [5].

Овај приступ се заснива на одвојеном моделовању случајних величина  $E[Y_i(1)|X_i]$  и  $E[Y_i(0)|X_i]$ , где за прву величину користимо само податке о кли-

<sup>1</sup>Слика преузета са <https://ambiata.com/blog/2020-07-07-uplift-modeling/>

јентима из групе са третманом, а за другу користимо податке о нетретираним клијентима. Предност овог приступа је његова једноставност. За решавање овог проблема могуће је користити класичне моделе машинског учења, независно од типа проблема (класификација или регресија). Оба модела могу остварити добре перформансе предвиђања одвојено, међутим за потребе моделовања уздизања, показано је да их често надјачавају друге методе [5], [32].

Један од разлога за то је што два модела одвојено предвиђају исходе, и зато могу да пропусте слабији ефекат уздизања. У контексту каузалне анализе, ефекат уздизања се односи на разлику између очекиваног исхода у третман групи и другој групи. Слабији ефекат значи да је разлика у очекиваном исходу између две групе мања, што значи да је утицај третмана мањи, тј. третман је мање ефикасан у побољшању исхода.

### 3.3 Приступ трансформације класа

Приступ трансформације класа је представљен у раду [16] у случају бинарне променљиве од значаја (енг. outcome variable)<sup>1</sup> ( $Y_i^{obs} \in \{0, 1\}$ ). Метода се састоји од креирања следеће циљне променљиве:

$$Z_i = Y_i^{obs}W_i + (1 - Y_i^{obs})(1 - W_i). \quad (3.1)$$

Нова циљна променљива има вредност 1 у следећим случајевима: посматрамо третирану групу и  $Y_i^{obs} = 1$  или посматрамо нетретирану групу и  $Y_i^{obs} = 0$ . У свим осталим случајевима, циљна променљива има вредност 0. Вероватноћа да је особа  $X_i$  у категорији особа подложних убеђивању (које желимо да третирамо) се записује као  $P(Z_i = 1|X_i)$ .

Дефинишемо вероватноћу да особа  $X_i$  добије третман

$$p(X_i) = P(W_i = 1|X_i).$$

Уводимо претпоставку о уравнотежености третиране и контролне групе у свим профилима појединца, која је први пут представљена у раду [16]. Ова претпоставка се односи на скуп свих варијабли тј. карактеристика које описују једног појединца, и подразумева да нема статистички значајних разлика

<sup>1</sup>Променљива од значаја се односи на променљиву која представља крајњи резултат у неком експерименту или студији. То је променљива коју желимо да анализирамо и схватимо шта узрокује њену појаву.

између ове две групе када се узму у обзир све карактеристике које описују појединца. Тада је  $p(X_i = x) = 1/2$  за све  $x$ , и показаћемо да важи следеће:

$$\tau(X_i) = 2P(Z_i = 1|X_i) - 1. \quad (3.2)$$

*Доказ.* Вероватноћа да особа припада категорији особа подложних убеђивању, записујемо на следећи начин:

$$P(Z_i = 1|X_i) = P(Y_i = 1, W_i = 1|X_i) + P(Y_i = 0, W_i = 0|X_i).$$

За сваку јединицу  $X_i$  можемо записати разлику између третираног и контролног исхода:

$$ITE = P(Y_i = 1|X_i, W_i = 1) - P(Y_i = 1|X_i, W_i = 0).$$

Како је:  $P(AB) = P(A|B)P(B)$ , вредност  $ITE$  можемо да напишемо на другачији начин.

$$ITE = \frac{P(Y_i = 1, W_i = 1|X_i)}{P(W_i = 1|X_i)} - \frac{P(Y_i = 1, W_i = 0|X_i)}{P(W_i = 0|X_i)}.$$

Вредности испод разломачке црте у оба случаја представљају вероватноћу да је особа у третманској групи односно у контролној групи. Претпоставка са почетка је да су третмани додељени насумично, што значи да су обе вероватноће једнаке  $\frac{1}{2}$ . Тако вредност  $ITE$  можемо записати:

$$ITE = 2[P(Y_i = 1, W_i = 1|X_i) - P(Y_i = 1, W_i = 0|X_i)].$$

Умањилац представља вероватноћу да особа  $X_i$  купи производ, и није у третираној групи. Ову вредност можемо да напишемо као разлику свих других случајева:

$$\begin{aligned} P(Y_i = 1, W_i = 0|X_i) &= 1 - P(Y_i = 1, W_i = 1|X_i) \\ &- P(Y_i = 0, W_i = 0|X_i) - P(Y_i = 0, W_i = 1|X_i). \end{aligned}$$

На основу тога, вредност  $ITE$  записујемо на следећи начин:

$$\begin{aligned} ITE &= 2[(P(Y_i = 1, W_i = 1|X_i) + P(Y_i = 0, W_i = 0|X_i)) \\ &+ \underbrace{(P(Y_i = 1, W_i = 1|X_i) + P(Y_i = 0, W_i = 1|X_i)) - 1}_{\text{све особе из третман групе}}] \end{aligned}$$

$$\begin{aligned}
&= 2[P(Z_i = 1|X_i) + P(W_i = 1|X_i) - 1] \\
&= 2\left[P(Z_i = 1|X_i) + \frac{1}{2} - 1\right].
\end{aligned}$$

Дакле,

$$ITE = 2P(Z_i = 1|X_i) - 1.$$

□

Моделовање уздицања се онда своди на моделовање  $P(Z_i = 1|X_i)$  (тј.  $E[Z_i = 1|X_i]$ ). Метода трансформације класа је заступљенија зато што обично има боље перформансе од описаног приступа два модела, а и даље задржава једноставност. Готово сваки класификатор може да се користи за моделирање  $E[Z_i = 1|X_i]$ .

У неким случајевима, полазне две претпоставке о бинарној резултујућој променљивој и уравнотеженим скуповима података, између контролне и третиране групе, могу деловати сувише рестриктивно. У раду [17], описана је генерализација неуравнотеженог додељивања третмана, где је предложено да се оцена *CATE* рачуна применом стандардних метода машинског учења на трансформисану циљну варијаблу  $Y_i^*$ . Кључно сазнање је да можемо описати условни просечни ефекат третмана као условно очекивање посматране променљиве путем трансформације исхода користећи индикатор третмана и вероватноћу доделе

$$Y_i^* = Y_i(1) \frac{W_i}{\hat{p}(X_i)} - Y_i(0) \frac{1 - W_i}{1 - \hat{p}(X_i)}, \quad (3.3)$$

где је величина  $\hat{p}(x)$  релевантна оцена оцене склоности припадања третманској групи (енг. propensity score-a)  $p(X_i)$ <sup>1</sup>. Кључно својство ове трансформисане варијабле је да у ситуацији где важе СИА услови, његово очекивање условљено  $X_i$  је једнако *CATE* [5], [2]:

$$E[Y_i^*|X_i] = \tau(X_i). \quad (3.4)$$

Ова особина значи да је свака релевантна оцена од  $E[Y_i^*|X_i]$ , такође релевантна оцена од  $\tau(X_i)$ .

У случају тоталне рандомизације ( $p(X_i = x) = 1/2$  за све  $x \in \{0, 1\}$ ) и бинарне циљне променљиве  $Y_i^{obs}$ , комбиновањем једначина (2.3) и (3.3),

<sup>1</sup>Када је величина узорка за тренинг велика, било који стандардни алгоритам машинског учења ће бити прикладан за оцену вероватноће  $p(X_i)$



можемо записати једначину (3.1) као:

$$Z_i = \frac{1}{2}Y_i^* + (1 - W_i), \quad (3.5)$$

и такође  $2E(Z_i|X_i) = E(Y_i^*|X_i) + 1$ , што је еквивалентно са једначином (3.2).

Претпоставка од које полазимо јесте претпоставка о независности (CIA)

$$W_i \perp (Y_i(0), Y_i(1)) \mid X_i.$$

Затим дефинишемо трансформисану величину

$$Y_i^* = Y_i^{\text{obs}} \cdot \frac{W_i - e(X_i)}{e(X_i) \cdot (1 - e(X_i))},$$

где је  $e(x) = P(W_i = 1 \mid X_i = x)$  условна вероватноћа третмана, тј. скор вероватноће. У случају потпуне рандомизације, скор вероватноће је константан  $e(x) = p$  за све  $x$ , а трансформација се поједностављује на:

$$Y_i^* = Y_i^{\text{obs}} \cdot \frac{W_i - p}{p \cdot (1 - p)},$$

где је  $p = \mathbb{E}[e(X_i)] = \mathbb{E}[W_i] = P(W_i = 1)$  заједничка вероватноћа доделе третману.

**Претпоставка 1.** Претпоставимо да важи горе поменути услов, тада је

$$\mathbb{E}[Y_i^* \mid X_i = x] = \tau(x).$$

*Доказ.* По дефиницији је

$$\begin{aligned} & \mathbb{E}[Y_i^* \mid X_i = x] \\ &= \mathbb{E}\left[Y_i^{\text{obs}} \cdot \frac{W_i - e(X_i)}{e(X_i) \cdot (1 - e(X_i))} \mid X_i = x\right] \\ &= \mathbb{E}\left[W_i \cdot Y_i^{\text{obs}} \cdot \frac{W_i - e(X_i)}{e(X_i) \cdot (1 - e(X_i))} \right. \\ & \quad \left. + (1 - W_i) \cdot Y_i^{\text{obs}} \cdot \frac{W_i - e(X_i)}{e(X_i) \cdot (1 - e(X_i))} \mid X_i = x\right]. \end{aligned}$$

Због  $W_i \cdot Y_i^{\text{obs}} = W_i \cdot Y_i(1)$  и  $(1 - W_i) \cdot Y_i^{\text{obs}} = (1 - W_i) \cdot Y_i(0)$ , записаћемо претходни израз на следећи начин:

$$\begin{aligned}
& \mathbb{E} \left[ W_i \cdot Y_i(1) \cdot \frac{W_i - e(X_i)}{e(X_i) \cdot (1 - e(X_i))} \right. \\
& \quad \left. + (1 - W_i) \cdot Y_i(0) \cdot \frac{W_i - e(X_i)}{e(X_i) \cdot (1 - e(X_i))} \mid X_i = x \right] \\
&= \mathbb{E} \left[ Y_i(1) \cdot \frac{W_i \cdot (1 - e(X_i))}{e(X_i) \cdot (1 - e(X_i))} \mid X_i = x \right] \\
& \quad - \mathbb{E} \left[ Y_i(0) \cdot \frac{(1 - W_i) \cdot e(X_i)}{e(X_i) \cdot (1 - e(X_i))} \mid X_i = x \right] \\
&= \mathbb{E} [Y_i(1) \cdot W_i \mid X_i = x] \cdot \frac{1}{e(X_i)} \\
& \quad - \mathbb{E} [Y_i(0) \cdot (1 - W_i) \mid X_i = x] \cdot \frac{1}{1 - e(X_i)}
\end{aligned}$$

Због претпоставке, ово је даље једнако:

$$\begin{aligned}
\mathbb{E} [Y_i^* \mid X_i = x] &= \mathbb{E} [Y_i(1) \mid X_i = x] \cdot \mathbb{E} [W_i \mid X_i = x] \cdot \frac{1}{e(X_i)} \\
& \quad - \mathbb{E} [Y_i(0) \mid X_i = x] \cdot \mathbb{E} [1 - W_i \mid X_i = x] \cdot \frac{1}{1 - e(X_i)} \\
&= \tau(x)
\end{aligned}$$

□

### 3.4 Приступ директног моделовања уздизања

Трећи и последњи приступ се састоји у модификовању постојећих алгоритама машинског учења како би се директно донео закључак о ефекту третмана. Различити модели су представљени у радовима, међу којима су: [2] за стратегију која се базира на логистичкој регресији, [30] и [12] који описују стратегију  $k$  најближих суседа, и [32] који показују модификацију методе потпорних вектора. Поред поменутих метода, најзаступљенији су модели засновани на стаблима одлучивања [14] и они који генерализују методе засноване на ансамблима [29]. У наставку ћемо описати методе засноване на стаблима одлучивања у случају бинарних стабала и дефинисаћемо критеријум поделе.

На примеру стабала одлучивања, модификација за потребе одређивања ефекта уздизања подразумева прилагођавање алгорита како би се директно предвиђао пораст ефекта уздизања. Уместо да дрво одлучивања предвиђа

циљну променљиву, оно се тренира тако да предвиђа ефекат третмана. То значи да алгоритам учи да разликује појединце који ће имати позитиван ефекат третмана од оних који неће имати никакав ефекат третмана или ће имати негативан ефекат третмана.

Постоји неколико метода за постизање ове модификације:

- **Методи на основу трансформације;**
- **Директни резултат пораста;**
- **Критеријуми за поделу:** У стаблу одлучивања, алгоритам бира најбољу поделу за чвор на основу одређених критеријума (нпр. Ђинијев коефицијент чистоће). У овој модификацији, критеријуми за поделу су прилагођени како би се оптимизовала предикција ефекта третмана уместо самог исхода.

У наставку ћемо описати методе засноване на стаблима одлучивања, у случају бинарних стабала и дефинисаћемо критеријум поделе.

Формално, у случају балансираног рандомизираниог експеримента, где је вероватноћа припадања третманској групи,  $p(X_i = x) = \frac{1}{2}$  за све  $x$ , оцена просечног ефекта третмана (или уздизања)  $\hat{\tau}$  је дата са:

$$\hat{\tau} = \underbrace{\frac{\sum_i Y_i^{obs} W_i}{\sum_i W_i}}_p - \underbrace{\frac{\sum_i Y_i^{obs} (1 - W_i)}{\sum_i (1 - W_i)}}_q. \quad (3.6)$$

Овај израз представља разлику просечног исхода између третираних и нетретираних клијената из обсервације.

Први критеријум поделе, описан је у раду [14], представља разлику у уздизању (позитивном утицају) између две гране:

$$\Delta = |\hat{\tau}_{Left} - \hat{\tau}_{Right}|, \quad (3.7)$$

где се индекси *Left*, *Right* односе на оцену из претходне једначине (3.6), коришћењем обсервација у листовима на левој и десној страни након поделе. Листови на левој и десној страни након поделе представљају подскупове посматраних података. Када се примени критеријум поделе, подаци се поделе на две групе - једна која испуњава услов дефинисан критеријумом и друга која га не испуњава.

Други критеријуми поделе су описани у раду [15] и предлажу се три критеријума поделе заснована на теорији информација у облику:

$$\Delta_{gain} = D_{aftersplit}(P^T, P^C) - D_{beforesplit}(P^T, P^C), \quad (3.8)$$

где је  $D$  мера дивергенције,  $P^T$  је вероватноћа исхода у третираној групи а  $P^C$  вероватноћа исхода у контролној групи. Сва три критеријума се ослањају на меру дивергенције између вероватноћа исхода у третираној и контролној групи, и служе за одређивање оптималне поделе у генерисању стабала одлучивања у контексту процене *SATE-a*. Критеријуми су дефинисани као функције различитих метрика дивергенције, укључујући Кулбакову дивергенцију, Еуклидско растојање и  $\chi^2$  дивергенцију:

$$KL(P : Q) = \sum_{k=Left, Right} p_k \log \frac{p_k}{q_k} \quad (3.9)$$

$$E(P : Q) = \sum_{k=Left, Right} (p_k - q_k)^2 \quad (3.10)$$

$$\chi^2(P : Q) = \sum_{k=Left, Right} \frac{(p_k - q_k)^2}{q_k}. \quad (3.11)$$

Критеријум мери колико би подела стабла смањила дивергенцију између  $P^T$  и  $P^C$ , а циљ је пронаћи ону која максимизира тај добитак. Вредност  $k$  указује на ком листу рачунамо вредности  $p, q$  који су дефинисани у једначини (3.6).

Остали критеријуми су предложени у склопу *модела каузалног стабла* у раду [5].

## Глава 4

# Евалуација

У овом одељку биће представљено како се правилно оцењују и евалуирају модели уздизања и како се изводе класичне метрике коришћењем ознака из економетрије. Оцењивање и евалуација ових модела се значајно разликује од традиционалног оцењивања модела машинског учења. У машинском учењу је стандард коришћење унакрсне валидације: одвајамо податке у скуп за тренирање и тестирање, учимо модел на тренинг подацима, предвиђамо циљну променљиву на тест подацима и упоређујемо их са стварним вредностима. У оцењивању модела уздизања, унакрсна валидација је и даље валидна идеја, али више нема стварне вредности са којом бисмо упоредили јер не можемо истовремено посматрати ефекат третмана и нетретмана на истој особи, па не постоји стандардна вредност коју бисмо користили за процену перформанси модела.

За илустрацију различитих метрика, користимо податке о клијентима банке, који се налазе на *GitHub* репозиторијуму [1]. Скуп података садржи информације о 10000 појединаца који су приближно равномерно подељени на третирани скуп (4953 појединца) и контролни скуп података (5047 појединаца). Детаљно објашњење колоне се налази у пропратном фајлу, на истом репозиторијуму [1]. Циљна вредност је индикатор  $\{0, 1\}$  која представља одлив (енг. churn) корисника из банке. Постоји 19 категоричких варијабли које описују кориснике и подела 80/20 у процесу учења и тестирања.

Корисници су унапред подељени између 4 описане категорије: корисници подложни убеђивању, сумњичави корисници, сигурно незаинтересовани корисници и сигурно заинтересовани корисници.

Идеја овог примера је да упоредимо колико ће модели ефикасно пред-

виђати ефекат уздизања међу већ унапред класификованим клијентима. У овом примеру постоји снажан негативни ефекат с обзиром на то да је 25 процената скупа података симулирано да има сумњичаво понашање. Сумњичави клијент је појединац који ће престати да користи услуге банке након изложености активној промоцији, али неће престати користити услуге ако није изложен промоцији. У овом случају, четвртина популације би негативно реаговала на циљану акцију.

Имплементирана су два од три приступа која су описана раније: приступ два модела и приступ трансформације класа. За приступ два модела су коришћена два модела са појачаним градијентом, док је приступ трансформације класа имплементиран путем алгоритма случајне шуме. Није коришћен метод решетке (енг. Grid-Search) за добијање најбољих хиперпараметара за ове моделе, јер је циљ овог примера био да се илуструју различите методе, а не да се утврди да ли је један бољи од другог.

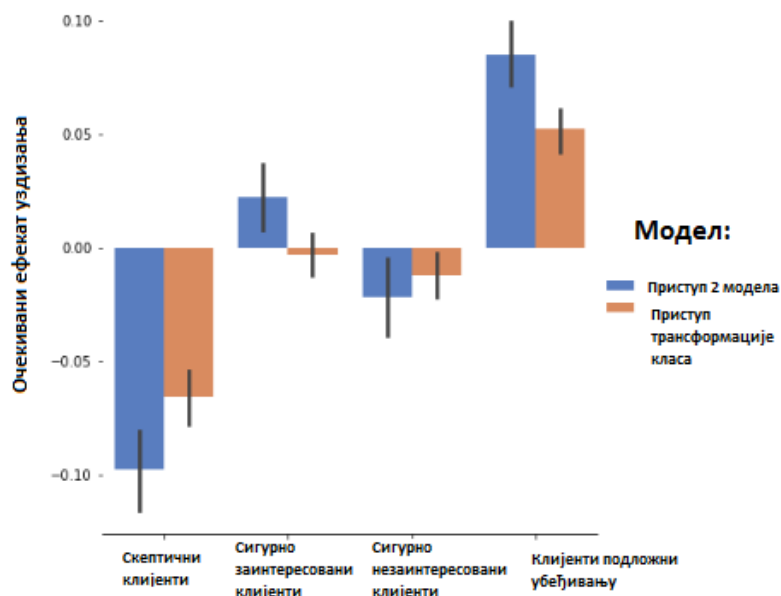
## 4.1 Традиционалне статистике уздизања

Традиционалне метрике коришћене за моделе уздизања су: Кини коефицијент, област испод криве уздизања, добијени прираст (енг. Uplift Gains), и неке од њих биће описане у наставку поглавља. Велики део литературе о моделима уздизања се бави проблемом који објашњава како није могуће истовремено посматрати и третирани и контролни исход за једну особу, што отежава израчунавање мере губитка (енг. loss measure) за свако посматрање. Као резултат, већина радова се ослања на агрегиране мере као што су *групе* ефекта уздизања (енг. uplift bins) [3] или *криве* ефекта уздизања (енг. uplift curves) [15], [16], [29], [17].

Правимо одвојено приступ два модела и трансформације класа. Коришћењем оба приступа добијамо предвиђене вредности ефекта уздизања за сваку особу у скупу. Савршен модел би имао вредност ефекта уздизања:

- 1: за категорију клијената подложних убеђивању;
- -1: за категорију сумњичавих клијената које не треба таргетирати;
- 0: за категорије клијената који неће променити своје акције независно од тога да ли добијају или не добијају третман (преостале две групе клијената).

Представићемо дијаграм приступа 2 модела и трансформације класа у оцењивању ефекта уздизања за сваку од 4 категорије клијената (Слика 4.1).



Слика 4.1: Оцена ефекта уздизања за све 4 категорије клијената

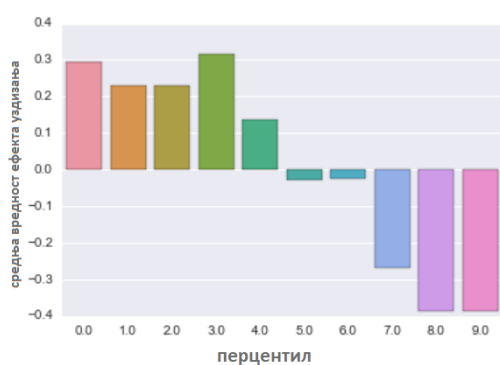
Посматрајући претходни график примећујемо да су вредности ефекта третмана ниске, што наводи на закључак да ови модели нису ефикасни. Разматрајући категорију клијената подложних убеђивању и сумњичавих клијената, приступ два модела има боље перформансе, док је за групе клијената код којих третман не би требало да има ефекта, приступ трансформације класа бољи.

Уобичајени приступ за евалуацију модела уздизања је да се прво предвиди ефекат уздизања за третиране и контролне опсервације, и да се израчуна просечно предвиђање по децилима<sup>1</sup> у обе групе. Затим се рачуна разлика између тих просека за сваки децил. Израчуната разлика нам даје идеју о добити уздизања по сваком децилу.

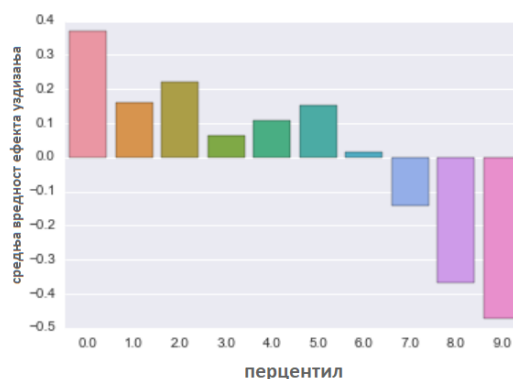
На слици 4.2, приступ два модела нам даје ефекат уздизања од 0.3 у првом децилу, док нам приступ класне трансформације даје ефекат уздизања од 0.4 (Слика 4.3). На слици је тешко ефективно упоредити моделе, и модел трансформације класа (Слика 4.3) се чини да боље функционише у првом

<sup>1</sup> Децили су статистичка подела узорка или скупа података на десет једнаких делова, тако да сваки део представља 10% укупног узорка. Децили се често користе за процену перформанси модела кроз различите сегменте популације.

децилу али не и у другом. У конкретном примеру, где циљна променљива представља индикатор да ли је клијент напустио банку, ефекат уздицања представљен на графицима указује да бисмо применом третмана спречили одлив клијената за одређени проценат по сваком децилу.

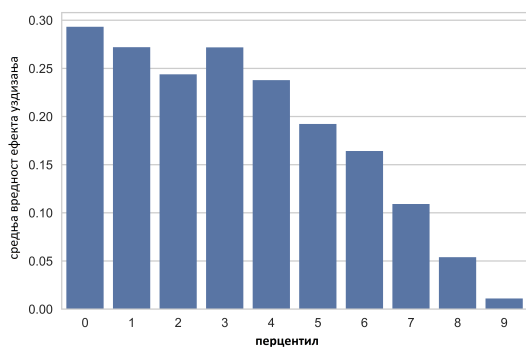


Слика 4.2: Добит уздицања за приступ 2 модела

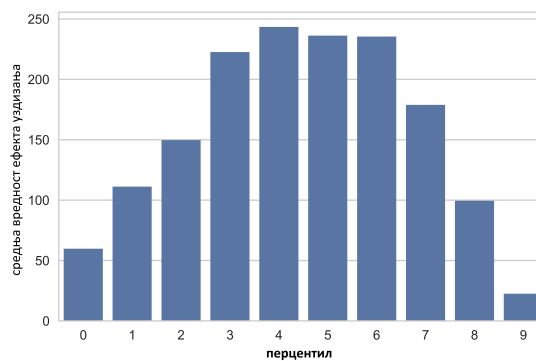


Слика 4.3: Добит уздицања за приступ трансформације класа

Да бисмо имали јаснију слику, приказаћемо кумулативне дијаграме децила на приступу два модела. На слици 4.4, прва трака са леве стране представља ефекат уздицања у првих 10 посто, следећа трака представља ефекат уздицања у првих 20 посто итд. Модел са добрим перформансама одликују високе вредности у првом децилу и опадајуће вредности у већим децилима.



Слика 4.4: Кумулативни дијаграм децила



Слика 4.5: Дијаграм кумулативне добити

На крају можемо погледати дијаграм кумулативне добити: рачунамо ефекат уздицања помножен са бројем особа које су узете у обзир (за сваку групу



енг. bin),

$$\left( \frac{Y^T}{N^T} - \frac{Y^C}{N^C} \right) (N^T + N^C), \quad (4.1)$$

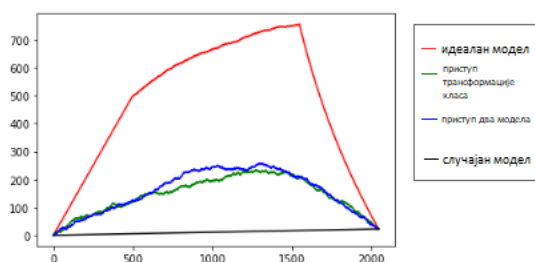
где су  $Y^T$  (односно  $Y^C$ ) збир исхода третираних појединаца (односно контролних) и  $N^T$  (односно  $N^C$ ) број третираних посматрања у групи (односно број контролних). Слика 4.5 приказује пример у случају приступа два модела. Овај график је користан за маркетаре јер се лако може уочити да ли третман има глобални позитиван или негативан ефекат и да ли се може очекивати боља добит циљањем дела популације. Тако можемо изабрати дециле који максимизирају добит као границу популације која ће бити третирана.

Проблем са графиконима је тај што не пружају никакву метрику и стога се не могу тачно упоређивати модели. Ипак, лако можемо генерализовати кумулативну криву укупног добитка (енг. cumulative gain chart) за сваку обсервацију тест скупа, са следећом параметарском кривом уздизања (енг. uplift curve) дефинисаном за свако  $t$  као:

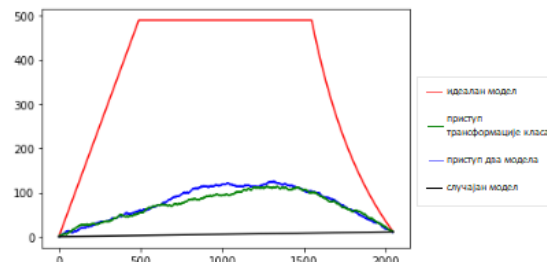
$$f(t) = \left( \frac{Y_t^T}{N_t^T} - \frac{Y_t^C}{N_t^C} \right) (N_t^T + N_t^C), \quad (4.2)$$

где индекс  $t$  указује на то да се количина израчунава у првих  $t$  посматрања, сортираних према израчунатој предвиђеној вредности уздизања.

На слици 4.6 је приказан пример таквих кривих и насумичне линије која одговара глобалном ефекту третмана. Благи позитивни нагиб ове насумичне линије значи да је третирање целе популације у целини корисно. Свака тачка на кривуљи одговара процењеној добити побољшања. Што је вредност виша, то је модел бољи. Континуираност криве уздизања чини могућим израчунавање површине испод криве као начина процене и поређења различитих модела уздизања. Ова мера је слична  $AUC$  површини испод  $ROC$  криве у бинарном класификационом скупу.



Слика 4.6: Криве уздизања



Слика 4.7: Кини криве

У нашем случају приступ два модела се чини бољи од приступа трансформације класа зато што је његова крива ближа црвеној линији која представља криву савршеног модела. Облик звона криве показује снажан позитиван и негативан ефекат присутан у скупу података.

У литератури, криве уздизања се често дефинишу као разлика између два утицаја израчуната на третираном и контролном скупу података. То можда није идеално јер немамо гаранцију да су најбоље рангирани примери у третираном и контролном скупу слични. Међутим, добро функционише у пракси и у случају насумичних и уравнотежених експеримената - две методе се поклапају/конвергирају.

Оригинална Кинијева (енг. Qini) крива је представљена у [22] као параметарска крива са следећом једначином:

$$g(t) = Y_t^T - \frac{Y_t^C N_t^C}{N_t^C}. \quad (4.3)$$

Кинијев коефицијент се дефинише као површина испод Кинијеве криве, који приказује однос између кумулативних позитивних реакција (одзива) и кумулативног броја јединица, сортираних по предикцијама у опадајућем редоследу. Висок Кини коефицијент, односно велика површина испод Кини криве, указује на боље перформансе модела у циљању циљне популације. Примери таквих кривих су дате на слици 4.7.

На истим графиконима (Слика 4.6 и 4.7) можемо упоредити и изглед кривих за идеалан и случајан модел. Крива случајног модела представљена је преко променљиве која је дефинисана као разлика просечних резултата за третманску групу и просечних резултата за контролну групу. Пошто је у питању случајни модел, третманска и контролна група се не разликују на основу било ког предиктивног модела, и разлика у резултатима је једноставно разлика између ове две просечне вредности. Идеалан модел је хипотетички

могућ али је у пракси његово постизање готово немогуће. У оваквом моделу, предвиђени резултати за таргетирану групу ће бити једнаки 1, а за контролну групу ће бити једнаки 0.

За оба приступа има места за побољшање јер смо у оба случаја далеко од црвене идеалне линије модела. Побољшања зависе од природе проблема и података, а неки од могућих су: прикупљање веће количине података, други одабир модела, бољи одабир и анализа варијабли које се користе у моделу.

## 4.2 Метрике базиране на $Y^*$

У поглављу 3 смо представили трансформацију циљне варијабли  $Y_i^*$ . Природно је запитати се да ли се функција ризика (нпр. средњеквадратна грешка  $MSE$ ), између наше оцене ( $\hat{\tau}$ ) и стварне вредности  $Y_i^*$ ,

$$MSE(Y_i^*, \hat{\tau}) = \sum_i^N \frac{1}{N} (Y_i^* - \hat{\tau}_i)^2 \quad (4.4)$$

може користити у схеми унакрсне валидације. Претходна једначина је апроксимација за

$$MSE(\tau_i, \hat{\tau}) = \sum_i^N \frac{1}{N} (\tau_i - \hat{\tau}_i)^2. \quad (4.5)$$

Иако се последња једначина може израчунати за симулиране податке где познајемо прави узрочни ефекат  $\tau_i$ , немогуће ју је извести из опсервацијских података [5]. Пошто се  $MSE$  не може израчунати, аутори уводе оцену која се може користити у стаблу одлучивања. Приступ оцене (енг. Estimation approach) је обавезан за оцену повећања ефекта уздизања. Предност коришћења ове метрике је у томе што не зависи од одабраног модела машинског учења.

Циљ је показати да можемо заменити  $Y_i^*$  са непознатим (енг. unobserved)  $\tau_i$  у  $MSE$  метрици како бисмо проценили перформансе нашег модела. Додавањем а затим одузимањем  $Y_i^*$  у претходној једначини, добијамо:

$$\begin{aligned} MSE &= \sum_i^N \frac{1}{N} (\tau_i - Y_i^* + Y_i^* - \hat{\tau}_i)^2 \\ &= \sum_i^N \frac{1}{N} [(\tau_i - Y_i^*)^2 + 2((\tau_i - Y_i^*)(Y_i^* - \hat{\tau}_i)) + (Y_i^* - \hat{\tau}_i)^2] \\ &\rightarrow_p E[(\tau_i - Y_i^*)^2] + 2E[(\tau_i - Y_i^*)(Y_i^* - \hat{\tau}_i)] + E[(Y_i^* - \hat{\tau}_i)^2]. \end{aligned}$$

Како бисмо минимизирали MSE, можемо занемарити први сабирак у једначини јер не зависи од вредности  $\hat{\tau}_i$ . У другом сабирку уочавамо да је  $\hat{\tau}_i \perp \tau_i | X_i$  и  $\hat{\tau}_i \perp Y_i^* | X_i$  јер је оцена базирана на функцији од  $X_i$ . Као резултат имамо:

$$\begin{aligned}
& E[(\tau_i - Y_i^*)(Y_i^* - \hat{\tau}_i)] \\
&= E[\tau_i Y_i^* - \tau_i \hat{\tau}_i - (Y_i^*)^2 + Y_i^* \hat{\tau}_i] \\
&= E[E[\tau_i Y_i^* - \tau_i \hat{\tau}_i - (Y_i^*)^2 + Y_i^* \hat{\tau}_i | X_i]] \\
&= E[E[\tau_i Y_i^* | X_i]] - E[E[\tau_i \hat{\tau}_i | X_i]] \\
&\quad - E[E[(Y_i^*)^2 | X_i]] + E[E[Y_i^* \hat{\tau}_i | X_i]] \\
&= E[E[\tau_i Y_i^* | X_i]] - \underbrace{E[\tau(X_i)E[\hat{\tau}_i | X_i]]}_{\tau_i \perp \hat{\tau}_i | X_i} \\
&\quad - E[E[(Y_i^*)^2 | X_i]] + \underbrace{E[\tau(X_i)E[\hat{\tau}_i | X_i]]}_{\hat{\tau}_i \perp Y_i^* | X_i, E[Y_i^* | X_i] = \tau(X_i)} \\
&= E[E[\tau_i Y_i^* | X_i]] - E[E[(Y_i^*)^2 | X_i]].
\end{aligned}$$

Како претходна једначина не зависи од оцено, видимо да асимптотски минимизирање оцено  $MSE(Y_i^*, \hat{\tau})$  доводи до минимизирања оцено  $MSE(\tau_i, \hat{\tau}_i)$ , што смо желели да покажемо.

## Глава 5

# Преглед библиотеке Causal ML и Meta-Learner алгоритама

Машинско учење и каузално закључивање су научне области које су се развијале независно једна од друге. Међутим постоји пресек између ова два поља. Causal ML је Python пакет који пружа сет метода за моделовање ефекта уздизања и каузално закључивање, користећи алгоритме машинског учења. Пружа кориснику стандардно окружење које нам омогућава да проценимо условне просечне ефекте третмана *CATE* или индивидуалне ефекте третмана *ITE* на основу екперименталних и посматраних података.

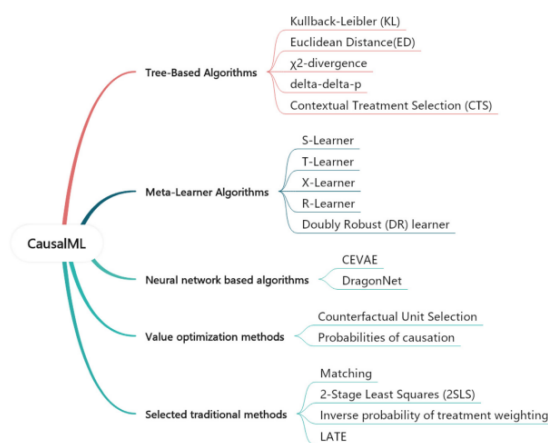


Fig. 2. Causal ML Algorithm Diagram.

Слика 5.1: Различити модели доступни у пакету

Тренутна верзија овог пакета пружа 15 алгоритама за моделовање ефекта

уздицања<sup>1</sup>. Мета алгоритми се ослањају на основни алгоритам како би проценили вредност просечног ефекта третмана (ATE).

## 5.1 Meta-Learner алгоритми

Мета-алгоритам или Meta-Learner је оквир за оцену условног просечног ефекта третмана (CATE) користећи било које алгоритме машинског учења. Мета алгоритам користи или појединачни модел при чему се индикатор третмана користи као варијабла (S-Learner), или више основних модела одвојено за третирану и контролну групу (T-Learner, X-Learner). S-Learner, T-Learner и X-Learner нису функције или библиотеке, већ различите врсте алгоритама које се користе у моделовању уздицања.

Поставка:

- Индексирамо сваког појединца са  $i$ .
- Сваки појединац је или у третманској ( $W_i = 1$ ) или у контролној групи ( $W_i = 0$ ).
- За сваког појединца имамо информације тј. предефинисане варијабле, коваријате  $X_i$ .
- Имамо одређену метрику исхода тј. одговора која нас занима. Ако је појединац у контролној групи (или у третманској), мера одговора је  $Y_i(0)$  (или  $Y_i(1)$ ). Као што смо поменули у претходном поглављу, имамо могућност да посматрамо само један од тих исхода, што обележавамо са  $Y_i^{obs} = Y_i(W_i)$ .

### S-Learner

S-Learner процењује ефекат третмана користећи један модел машинског учења, користећи варијаблу третмана  $W_i$  као оне у вектору  $X_i$ .

- Корак 1: Процењујемо просечан исход  $\widehat{\mu(x, w)}$  са коваријатама  $X$  и варијаблом индикатора третмана  $W$  користећи модел машинског учења,

<sup>1</sup>Слика је преузета са сајта <https://www.sciencedirect.com/science/article/pii/S2352711022002126>

где израз

$$\mu(x, w) = E[Y^{obs} | X = x, W = w]$$

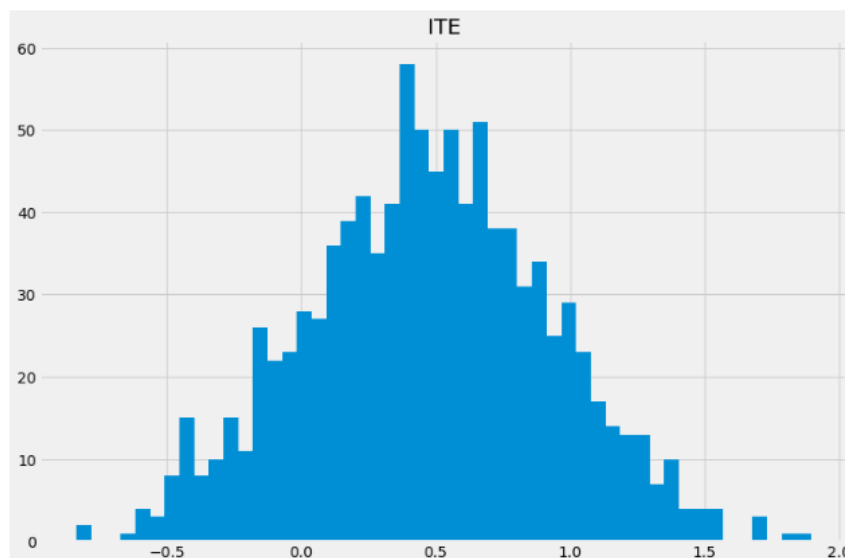
представља модел S-Learnera.

- Корак 2: Оцењујемо вредност *CATE*:

$$\hat{\tau}(x) = \hat{\mu}(x, W = 1) - \hat{\mu}(x, W = 0).$$

Пример коришћења овог модела на синтетичким подацима приказан је на *GitHub* репозиторијуму [1]. На истом репозиторијуму се могу наћи идентични примери и за друге две методе, али их у раду нећемо приказати. Подаци су креирани коришћењем методе из *Causal ML* пакета, где су клијенти унапред распоређени у третманску и контролну групу. Циљ примера је да предвиди индивидуални ефекат третмана за сваког клијента и да се одлучи којим клијентима ће бити послата промотивна понуда.

Хистограм индивидуалног ефекта третмана показује нормалну расподелу (Слика 5.2), где је просечна вредност око 0.5 и да узорци углавном имају позитиван ефекат третмана. Просечна вредност ефекта третмана која је добијена коришћењем модела је 0.5.



Слика 5.2: Расподела вредности индивидуалног ефекта третмана

Важно је дефинисати вредност прага и на тај начин добро распоредити клијенте у 3 категорије: клијенти подложни убеђивању (њих желимо да таргетирамо), сумњичави клијенти који ће куповати производ само ако их не

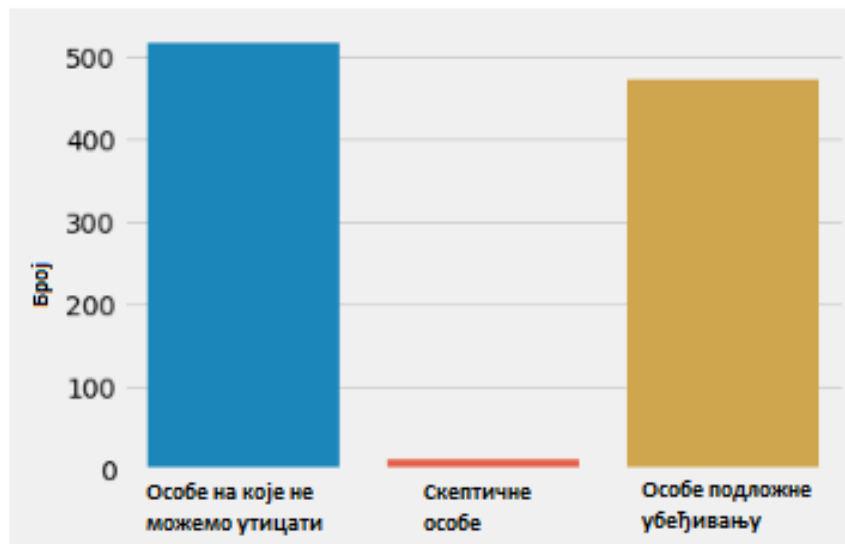
контактирамо и клијенти без утицаја, независно од наше акције они ће куповати или неће куповати, тако да на њих нећемо трошити ресурсе.

```

1 # Create the segmentation function
2 def segmentation(x):
3     if x < -0.5:
4         return 'sleeping-dog'
5     elif x > 0.5:
6         return 'persuadable'
7     else:
8         return 'no_impact'
9
10 # Apply the function
11 ite_df['segment'] = ite_df['ITE'].apply(segmentation)
12
13 # Visualization
14 sns.countplot(x=ite_df['segment'])

```

На слици 5.3 је приказана класификација клијената у 3 категорије које су нам од значаја на основу резултата модела. На овај начин можемо сазнати који су то клијенти који су сумњичави и на које не можемо извршити никакав утицај, али и ко су клијенти подложни убеђивању које ћемо имати у виду за будуће контактирање.



Слика 5.3: Сегментација клијената на основу резултата модела



## T-Learner

T-Learner се састоји од 2 корака и користи 2 модела:

- Корак 1: Користимо посматрања у контролној групи да бисмо проценили одговор у контролној групи  $\hat{\mu}_0(x)$ , за  $\mu_0(x) = E[Y(0)|X = x]$ . Слично томе, користимо посматрања из третманске групе да бисмо проценили одговор у тој групи  $\hat{\mu}_1(x)$ , за  $\mu_1(x) = E[Y(1)|X = x]$ . Користимо било који модел машинског учења.
- Корак 2: Оцењујемо вредност *CATE*:

$$\hat{\tau}(x) = \hat{\mu}_1(x) - \hat{\mu}_0(x).$$

## X-Learner

Уводимо још један алгоритам, X-Learner, који представља побољшање претходно описаног T-Learner алгоритма, и састоји се из 3 корака.

- Корак 1: Користимо посматрања у контролној групи да бисмо проценили одговор у контролној групи  $\hat{\mu}_0(x)$ , за  $\mu_0(x) = E[Y(0)|X = x]$ . Слично томе, користимо посматрања из третманске групе да бисмо проценили одговор у тој групи  $\hat{\mu}_1(x)$ , за  $\mu_1(x) = E[Y(1)|X = x]$ . Користимо било који модел машинског учења.
- Корак 2: Користимо процене из првог корака да бисмо добили процене индивидуалних ефеката третмана (*ITE*). За опсервације у контролној групи, процена *ITE* се рачуна као  $\hat{D}_i = \hat{\mu}_1(X_i) - Y_i^{obs}$ ; За опсервације у третманској групи, процена *ITE* се рачуна као  $\hat{D}_i = Y_i^{obs} - \hat{\mu}_0(X_i)$ . Правимо модел за *ITE* користећи само опсервације из контролне групе (са импутираним/процењеним *ITE* као одговором),  $\hat{\tau}_0(x)$ . Сличну калкулацију урадимо за опсервације из третманске групе како бисмо добили  $\hat{\tau}_1(x)$ .
- Корак 3: Оцењујемо вредност *CATE* комбинујући претходно добијене вредности, користећи и функцију тежине  $g \in [0, 1]$ . Дobar одабир за ову функцију је припадност третманској групи (скор склоности):

$$\hat{\tau}_X(x) = g(x)\hat{\tau}_0(x) + [1 - g(x)]\hat{\tau}_1(x).$$

Поставља се питање када ћемо користити који алгоритам:

- Избор основног модела може значајно утицати на тачност предикције. Ово је важна предност мета-алгоритама уопште. Не постоји универзално најбољи мета-алгоритам, за сваки од ова три мета-алгорита постоје ситуације у којима се најбоље показују.
- **T-Learner:** Добро се показују ако нема заједничких трендова у одговору у контролној групи и одговору у третираној групи, као и ако је ефекат третмана врло комплексан. Због тога што се подаци не обједињују преко третманских група, T-Learner-у је тешко да имитира понашање (на пример, дисконтинуитет) који се јавља у свим третманским групама.
- **S-Learner:** Пошто индикатор третмана не игра посебну улогу, основни модели га могу потпуно занемарити приликом прилагођавања модела. Ово је добро ако је *CATE* нула на многим местима.
- **X-Learner:** X-Learner може се прилагодити структурним својствима као што су реткост или глаткоћа *CATE*-а (ово је корисно јер је *CATE* често нула или приближно линеарна). Када је *CATE* нула, обично није једнако добар као S-Learner, али је бољи од T-Learnerа. Када је *CATE* комплексан, X-Learner даје боље резултате од S-Learnerа, и често је бољи и од T-Learnerа. Посебно је ефикасан када је број јединица у једној третманској групи (често контролној групи) знатно већи од броја јединица у другој групи.

## Глава 6

# Примена модела уздизања на реалне податке

У овом поглављу биће представљена примена модела уздизања и каузалног закључивања на реалне податке. Показаћемо примену модела који су појашњени у претходном поглављу (*Meta-Models*) уз помоћ библиотека *CausalML* и *scikit-uplift*.

Примена модела ће бити приказана на примеру „Програма лојалности банке”. Програм лојалности банке је стратегија коју банка имплементира како би награђивала и задржавала постојеће клијенте. Овај програм има за циљ подстицање клијената да редовно користе услуге банке, као и да се повежу са различитим производима које банка нуди. То може укључивати попусте на камате, ниже накнаде за трансакције, повећане лимите кредитних картица, бонус поене итд. Платформа нуди широк спектар награда на које клијенти могу искористити своје бодове, а алтернативно се бодови могу искористити и за поврат новца.

Кодови су писани у програмском језику *Python* и могу се пронаћи на следећем *GitHub* репозиторијуму [1].

Скуп података садржи неколико табела са подацима који нам дају различите информације о клијентима банке и може се пронаћи на: [27]. За сваког клијента из базе, идеја је да се оцени ефективност комуникације упућене њима тј ефекат уздизања. Клијенти из скупа представљају особе које ће банка бирати, да ли ће их контактирати за потребе програма лојалности, који за циљ има повећање продаје одређених производа.

У питању су следећи скупови података:

- *clients.csv* - Опште информације о клијентима;
- *purchases.csv* - Историја коришћења услуга клијента пре комуникације са њим;
- *uplift\_train.csv* - Подскуп клијената за тренирање модела; Две најбитније варијабле у овом скупу су *treatment\_flg* и представља информацију да ли је клијент био контактиран или није; и варијабла *target* која показује да ли је клијент купио неки од производа након контактирања;
- *uplift\_test.csv* - Подскуп клијената за тестирање на ком ћемо оцењивати вредност ефекта уздизања.

Клијенти су унапред подељени у третманску и контролну групу водећи рачуна о претпоставкама које су описане у првом поглављу. У неком од претходних поглавља представили смо поделу клијената у 4 категорије. Приликом контактирања клијената за потребе кампање, желимо да повећамо продају тј. профит, али и да избегнемо непотребне трошкове приликом контактирања. У складу са тим, не желимо да контактирамо све клијенте већ да допремо до оних који су значајни.

Постоје 3 могућа исхода када одлучимо да ли треба да контактирамо особу, и подела је на основу историјских података:

- +1: Клијент је купио производ након што смо га контактирали и вероватно не би обавио куповину да није био контактиран. По подели ово су особе подложне убеђивању.
- 0: Без обзира на то што смо се обратили овом клијенту, он би обавио куповину (или је не би обавио, у овом случају је то подједнако значајно). По подели ово су сигурно заинтересовани или незаинтересовани клијенти.
- -1: Клијент није купио производ зато што смо га контактирали, али да нисмо, он би купио производ. По подели ово су скептични клијенти.

## 6.1 Упознавање са подацима

Табела *clients.csv* садржи податке о 400 хиљада клијената као што су: јединствени број тј. *ID*, године, пол, датум првог пријављивања на платформу и датум када је клијент користио неку од могућих услуга банке. Подаци су унапред подељени методом рандомизације на тренинг и тест скуп, тако да ћемо приказати расподелу клијената у тренинг скупу.

Као што је поменуто, варијабла *treatment* је бинарна варијабла која садржи вредности  $\{0, 1\}$  и представља индикатор да ли је клијент примио третман (тј. био контактиран) или није примио третман (није био контактиран, у контролној је групи).

Варијабла *target* такође представља бинарну варијаблу са вредностима  $\{0, 1\}$  и представља индикатор да ли је клијент након нашег контактирања купио производ тј. услугу.

Расподела циљне променљиве у односу на третирану и контролну групу је приказана у табели 6.1.

Таргет	Третманска група	Контролна група
0	17%	20%
1	33%	30%

Табела 6.1: Удео расподела циљне променљиве у односу на третирану и контролну групу

Имамо приближно једнаку поделу клијената у третираној и контролној групи.

Још једна од карактеристика које имамо за клијенте јесте категоричка варијабла пол (мушкарци, жене, неизјашњени). Избацујемо неке аутлајере у вези са овом варијаблом, и клијенте који имају непотпуне информације о датуму пријаве на платформу.

Како бисмо унапредили моделе које ћемо касније користити, израчунаћемо неке једноставне сумарне метрике за наш скуп података и сачувати их.

## 6.2 Припрема података

Од постојећих варијабли правимо додатне варијабле које ћемо користити за моделе уздицања. Категоричке варијабле трансформишемо у нумеричке коришћењем енкодера, и тренинг скуп података делимо на скуп за тренинг и валидацију у размери 70:30.

## 6.3 Модел склоности

Модел скора склоности, се дефинише као  $p(X_i) = P(W_i = 1|X_i)$ , што представља вероватноћу клијента да прими третман под његовим датим карактеристикама. Уколико је додела третмана насумична, ови резултати би требало да буду концентрисани око 0.5.

```

1 propensity_model = HistGradientBoostingClassifier(**hgc_params)
2
3 propensity_model.fit(X=x_train_transformed, y=w_train)
4 p_train = propensity_model.predict_proba(X=x_train_transformed)
5 p_val = propensity_model.predict_proba(X=x_val_transformed)
6
7 p_train = pd.Series(p_train[:, 0], index=idx_train)
8 p_val = pd.Series(p_val[:, 0], index=idx_val)

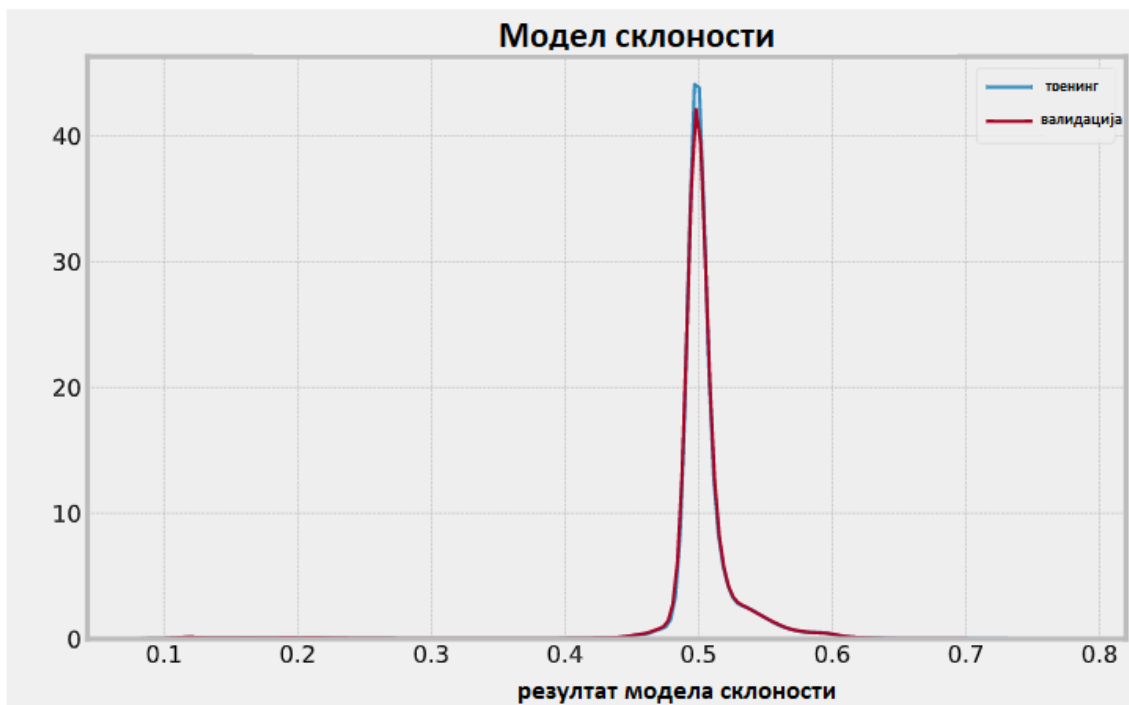
```

Представљамо график резултата на слици 6.3 где смо очекивали да је највећем броју клијената вредност појединачног ефекта третмана око 0.5:

```

1 fig, ax = plt.subplots()
2 sns.kdeplot(x=p_train, label="train", ax=ax)
3 sns.kdeplot(x=p_val, label="val", ax=ax)
4 ax.legend()
5 ax.set(
6     title="Propensity Score Predictions Distribution",
7     xlabel="propensity score",
8     ylabel="density",
9 );

```



Слика 6.1: Резултат модела склоности на тренинг скупу и скупу за валидацију

## 6.4 Модели

Након што смо представили податке са којима радимо, прелазимо на моделовање. Из претходно описане библиотеке *Causal ML*, користимо 3 различита модела: *S-Learner*, *T-Learner*, *X-Learner*.

### S-Learner

У овом моделу, тренирамо модел машинског учења (енг. Gradient Boosting класификатор) који предвиђа циљну променљиву користећи коваријате  $X_i$ , за  $\mu(x, w) = E[Y|X = x, W = w]$ .

- **Корак 1: Тренинг**

$$\underbrace{\begin{pmatrix} x_{1,1} & \cdots & x_{1,k} & w_1 \\ \vdots & \ddots & \vdots & \vdots \\ x_{n,1} & \cdots & x_{n,k} & w_n \end{pmatrix}}_{X \oplus W} \xrightarrow{\mu} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

• **Корак 2: Предвиђање ефекта уздизања**

$$\widehat{uplift} = \hat{\mu} \begin{pmatrix} x_{1,1} & \cdots & x_{1,k} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ x_{m,1} & \cdots & x_{m,k} & 1 \end{pmatrix} - \hat{\mu} \begin{pmatrix} x_{1,1} & \cdots & x_{1,k} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ x_{m,1} & \cdots & x_{m,k} & 0 \end{pmatrix}$$

```

1 s_learner = BaseSClassifier(learner=HistGradientBoostingClassifier(**
   hgc_params))
2
3 s_ate = s_learner.estimate_ate(
4     X=data_train.x_transformed, treatment=data_train.treatment, y=
   data_train.y
5 )

```

**T-Learner**

У овом моделу, тренирамо 2 модела машинског учења, један за третирану групу а други за контролну групу.

• **Корак 1: Тренинг**

$$\underbrace{\begin{pmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & \ddots & \vdots \\ x_{n_C,1} & \cdots & x_{n_C,k} \end{pmatrix}}_{X|_{control}} \xrightarrow{\mu_C} \begin{pmatrix} y_1 \\ \vdots \\ y_{n_C} \end{pmatrix}$$

$$\underbrace{\begin{pmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & \ddots & \vdots \\ x_{n_T,1} & \cdots & x_{n_T,k} \end{pmatrix}}_{X|_{treatment}} \xrightarrow{\mu_T} \begin{pmatrix} y_1 \\ \vdots \\ y_{n_T} \end{pmatrix}$$

• **Корак 2: Предвиђање ефекта уздизања**

$$\widehat{uplift} = \hat{\mu}_T \begin{pmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m,k} \end{pmatrix} - \hat{\mu}_C \begin{pmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m,k} \end{pmatrix}$$



```

1 t_learner = BaseTClassifier(learner=HistGradientBoostingClassifier(**
    hgc_params))
2
3 t_ate_lwr, t_ate, t_ate_upr = t_learner.estimate_ate(
4     X=data_train.x_transformed, treatment=data_train.treatment, y=
    data_train.y
5 )

```

## X-Learner

Овај модел је сличан претходном моделу, али укључује додатни корак где прослеђује информацију са једног на други модел.

- **Корак 1: Тренинг**

$$\underbrace{\begin{pmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & \ddots & \vdots \\ x_{n_C,1} & \cdots & x_{n_C,k} \end{pmatrix}}_{X|_{control}} \xrightarrow{\mu_C} \begin{pmatrix} y_1 \\ \vdots \\ y_{n_C} \end{pmatrix}$$

$$\underbrace{\begin{pmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & \ddots & \vdots \\ x_{n_T,1} & \cdots & x_{n_T,k} \end{pmatrix}}_{X|_{treatment}} \xrightarrow{\mu_T} \begin{pmatrix} y_1 \\ \vdots \\ y_{n_T} \end{pmatrix}$$

- **Корак 2: Рачунање импутованог ефекта третмана**

$$\tilde{D}^T := \begin{pmatrix} y_1 \\ \vdots \\ y_{n_T} \end{pmatrix} - \hat{\mu}_C \begin{pmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & \ddots & \vdots \\ x_{n_T,1} & \cdots & x_{n_T,k} \end{pmatrix}$$

$$\tilde{D}^C := \hat{\mu}_T \begin{pmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & \ddots & \vdots \\ x_{n_C,1} & \cdots & x_{n_C,k} \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_{n_C} \end{pmatrix}$$

- **Корак 3: Тренирање ефекта уздицања са различитим варија-  
блама**

$$\underbrace{\begin{pmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & \ddots & \vdots \\ x_{n_C,1} & \cdots & x_{n_C,k} \end{pmatrix}}_{X|_{control}} \xrightarrow{\tau_C} \begin{pmatrix} \tilde{D}_1^C \\ \vdots \\ \tilde{D}_{n_C}^C \end{pmatrix}$$

$$\underbrace{\begin{pmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & \ddots & \vdots \\ x_{n_T,1} & \cdots & x_{n_T,k} \end{pmatrix}}_{X|_{treatment}} \xrightarrow{\tau_T} \begin{pmatrix} \tilde{D}_1^T \\ \vdots \\ \tilde{D}_{n_T}^T \end{pmatrix}$$

• **Корак 4: Предвиђање ефекта уздизања**

$$uplift = g(x)\hat{\tau}_C(x) + (1 - g(x))\hat{\tau}_T(x),$$

где је  $g(x) \in [0, 1]$  тежинска функција.

Ефекат уздизања  $\tau$  се процењује са обе величине  $\hat{\tau}_C$  и  $\hat{\tau}_T$ , док се функција  $g$  бира тако да комбинује ове две оцене у једну унапређену оцену  $\hat{\tau}$ . Добро је користити оцену вероватноће припадности за функцију  $g$ , али такође има смисла одабрати и  $g = 1$  или  $g = 0$  ако је број третираних јединица веома велики или мали у поређењу са бројем контролних јединица.

```

1 x_learner = BaseXClassifier(
2     outcome_learner=HistGradientBoostingClassifier(**hgc_params),
3     effect_learner=HistGradientBoostingRegressor(**hgc_params),
4 )
5
6 x_ate_lwr, x_ate, x_ate_upr = x_learner.estimate_ate(
7     X=data_train.x_transformed,
8     treatment=data_train.treatment,
9     y=data_train.y,
10    p=data_train.p,
11 )

```

Ради упоређивања наша 3 модела, приказаћемо оцену просечног ефекта третмана на тест скупу. Као што је раније објашњено, ова величина представља оцену колико се просечно мења циљна променљива као резултат примене одређеног третмана у односу на контролну групу која није подвргнута третману. У нашем случају, циљ третмана је да се постигне што већи позитиван

утицај на циљну променљиву, тако да је већи  $ATE$  пожељан. У табели 6.2 видимо резултат за сва три алгорита, и примећујемо да T-Learner има најбољи резултат.

За најбољи модел, израчуната вредност од 0.02 указује на то да је модел постигао ефекат третмана од 2% у просеку у циљној променљивој у поређењу са контролном групом.

Модел	Просечан ефекат третмана (ATE)
S-Learner	0.0173
T-Learner	0.0204
X-Learner	0.0197

Табела 6.2: Просечан ефекат третмана за различите моделе

## 6.5 Предвиђање модела

Након што смо обучили наша 3 модела, урадићемо предикције на тренинг и валидационом скупу. То ћемо урадити за сва 3 модела одвојено.

```

1 #primer za S learner
2 class DataOut:
3     meta_learner_name: str
4     meta_learner: BaseLearner
5     y_pred: np.array
6
7
8 # in-sample predictions
9 data_out_train_s = DataOut(
10     meta_learner_name="S-Learner",
11     meta_learner=s_learner,
12     y_pred=s_learner.predict(
13         X=data_train.x_transformed, treatment=data_train.treatment
14     ),
15 )
16
17 # out-of-sample predictions
18 data_out_val_s = DataOut(
19     meta_learner_name="S-Learner",
20     meta_learner=s_learner,
21     y_pred=s_learner.predict(

```

```

22     X=data_val.x_transformed, treatment=data_val.treatment
23     ),
24 )

```

Природно питање које се овде поставља је колико добро можемо предвидети ефекат уздизања? Код моделовања ефекта побољшања, не постоји „стварна вредност” јер никада не можемо посматрати ефекат третмана и контроле (нетретмана) на истој особи истовремено. Да бисмо одговорили на ово питање, детаљније анализирамо неке дијагностичке алате модела.

## Ефекат уздизања по децилу

1. Сортирамо предвиђања ефекта уздизања у опадајућем редоследу.
2. Предвиђамо ефекте уздизања за третирану и контролну групу клијената.
3. Израчунамо просечно предвиђање по децилима у обе групе.
4. Разлика између тих просека се узима за сваки перцентил.

Ова разлика даје представу о добијеном ефекту уздизања по децилима. Ову функцију користимо из поменутог пакета *scikit.metrics*.

```

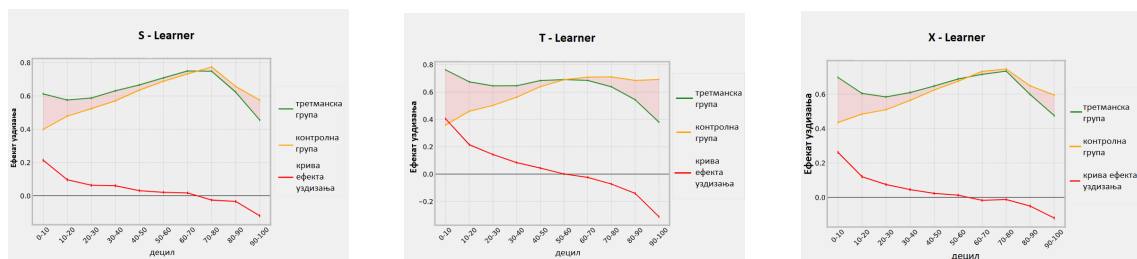
1
2 uplift_by_percentile_df = uplift_by_percentile(
3     y_true=data_train.y,
4     uplift=data_out_train.s.y_pred.flatten(),
5     treatment=data_train.treatment,
6     strategy="overall",
7     total=True,
8 )
9
10 uplift_by_percentile_df

```

	n_treatment	n_control	response_rate_treatment	response_rate_control	uplift
percentile					
0-10	7090	6862	0.612412	0.398863	0.213549
10-20	7024	6928	0.575456	0.478637	0.096818
20-30	7037	6915	0.587040	0.523644	0.063396
30-40	7013	6938	0.630543	0.570193	0.060350
40-50	7089	6862	0.665961	0.634946	0.031015
50-60	7019	6932	0.708790	0.688113	0.020677
60-70	7087	6864	0.749683	0.732226	0.017456
70-80	6910	7041	0.748625	0.774180	-0.025555
80-90	6870	7081	0.623144	0.656687	-0.033543
90-100	6583	7368	0.454960	0.575054	-0.120095
total	69722	69791	0.636743	0.603531	0.033213

Слика 6.2: Вредност ефекта уздицања по перцентилу

Добар модел би требало да има веће вредности у првим децилима и опадајуће вредности за веће дециле. Приказаћемо графике ефекта уздицања по децилима за сва 3 модела на тренинг скупу (Слика 6.3) и валидационом скупу (Слика 6.4).



Слика 6.3: Тренинг скуп: S-Learner (лево), T-Learner (средина) и X-Learner (десно)

На поменутиим графицима, зелена линија представља вредности предвиђеног ефекта уздицања на третираној групи, жута линија представља вредности предвиђеног ефекта уздицања на контролној групи и црвена линија представља ефекат уздицања који се рачуна као разлика просека претходне две вредности.



Слика 6.4: Валидациони скуп: S-Learner (лево), T-Learner (средина) и X-Learner (десно)

Посматрајући изглед графика ефекта уздизања тј. црвену линију, закључујемо да T-Learner и X-Learner имају сличне перформансе, и боље него први модел.

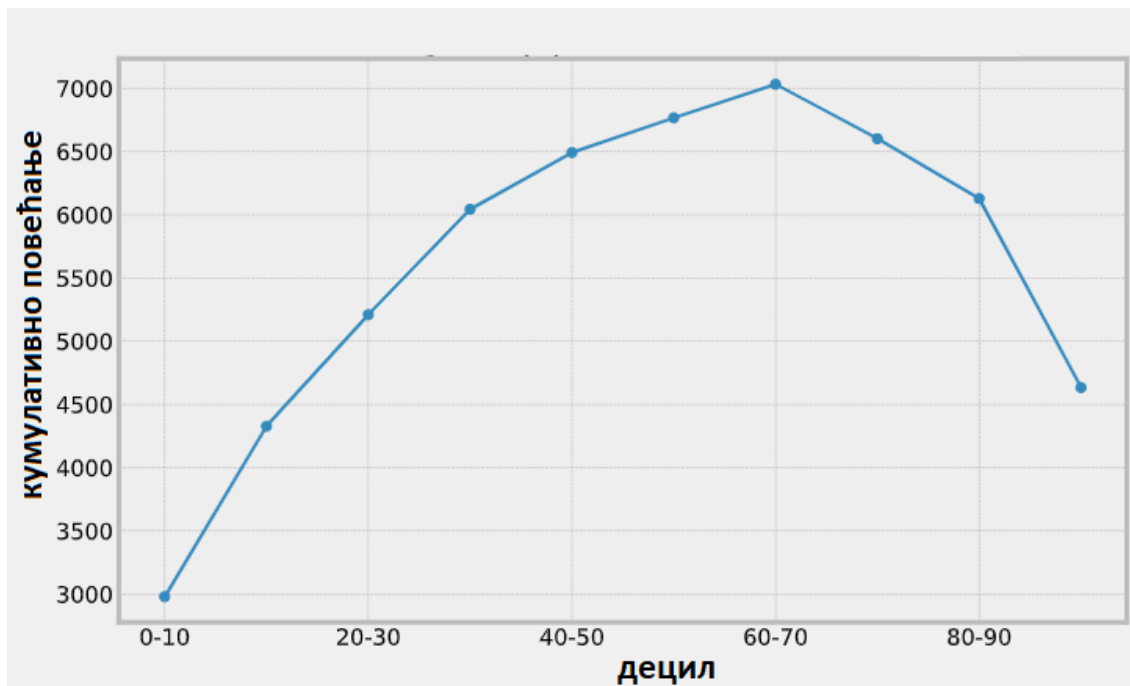
## Графикон кумулативног повећања

Предвиђамо ефекат уздизања за обе посматране групе - третирану и контролну, и рачунамо просечну предикцију по децилу (bin) у обе групе. Затим за сваки децил, рачунамо разлику између тих просека:

$$\left( \frac{Y^T}{N^T} - \frac{Y^C}{N^C} \right) (N^T + N^C),$$

- $Y^T/Y^C$ : Збир резултата третираних/контролних јединки у децилу.
- $N^T/N^C$ : Број третираних/контролних јединки у децилу.

Овај графикон је користан за тржишне стручњаке јер могу лако видети да ли третман има глобално позитиван или негативан ефекат и да ли могу очекивати бољи добитак циљајући део популације. Можемо тако изабрати децил који максимизује добитак као границу популације која ће бити циљана. Приказаћемо га на примеру S-Learner модела (Слика 6.5).



Слика 6.5: График кумулативног повећања на примеру S-Learner-а

Посматрајући графикон кумулативне добити, видимо да је вредност кумулативне добити позитивна (последња тачка на графику је вредност већа од 0). Та вредност означава да уколико би се промотивна понуда банке понудила свим клијентима, свакако бисмо имали позитиван ефекат уздицања. Међутим нама је у интересу да то оптимизујемо.

Такође можемо приметити да је вредност криве највећа око 60-70. перцентила, што нам на конкретном скупу података говори да је ефекат уздицања већи уколико бисмо у обзир узели мањи скуп клијената а не цео скуп. На тај начин бисмо уштедели ресурсе и повећали ефекат уздицања.

У пракси су нам ове информације од значаја када банка шаље поново промотивне понуде клијентима, где је најбоље направити пресек како би се максимизирала успешност контактирања и смањили ресурси, сходно томе да радимо са готово истим скупом података.

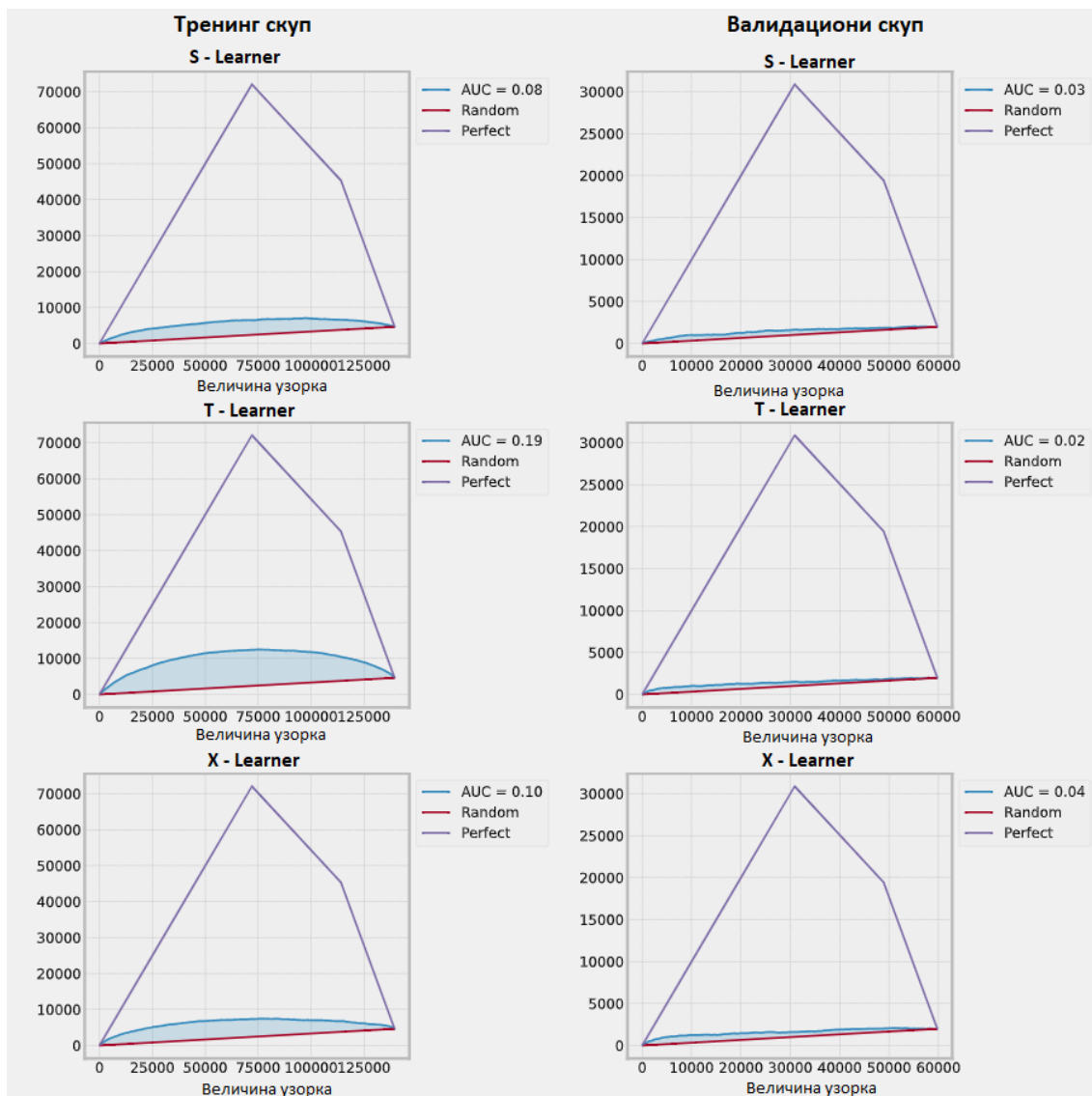
## Крива ефекта уздицања

Можемо генерализовати криву кумулативног повећања за свако посматрање тестног скупа, где индекс  $t$  означава да се рачуна првих  $t$  посматрања,

сортираних по претпостављеним вредностима ефекта побољшања,

$$f(t) = \left( \frac{Y_t^T}{N_t^T} - \frac{Y_t^C}{N_t^C} \right) (N_t^T + N_t^C).$$

Приказаћемо график ове криве за сва 3 модела, на тренинг и валидационом скупу (Слика 6.6).



Слика 6.6: Крива ефекта уздицања за сва 3 модела

На истом графику уочавамо и криву ефекта уздицања у случају савршеног модела (љубичаста линија) и у случају насумичног модела (црвена линија).

Савршен модел додељује више бодова свим третираним појединцима са позитивним резултатима него било ком појединцу са негативним резултатима.



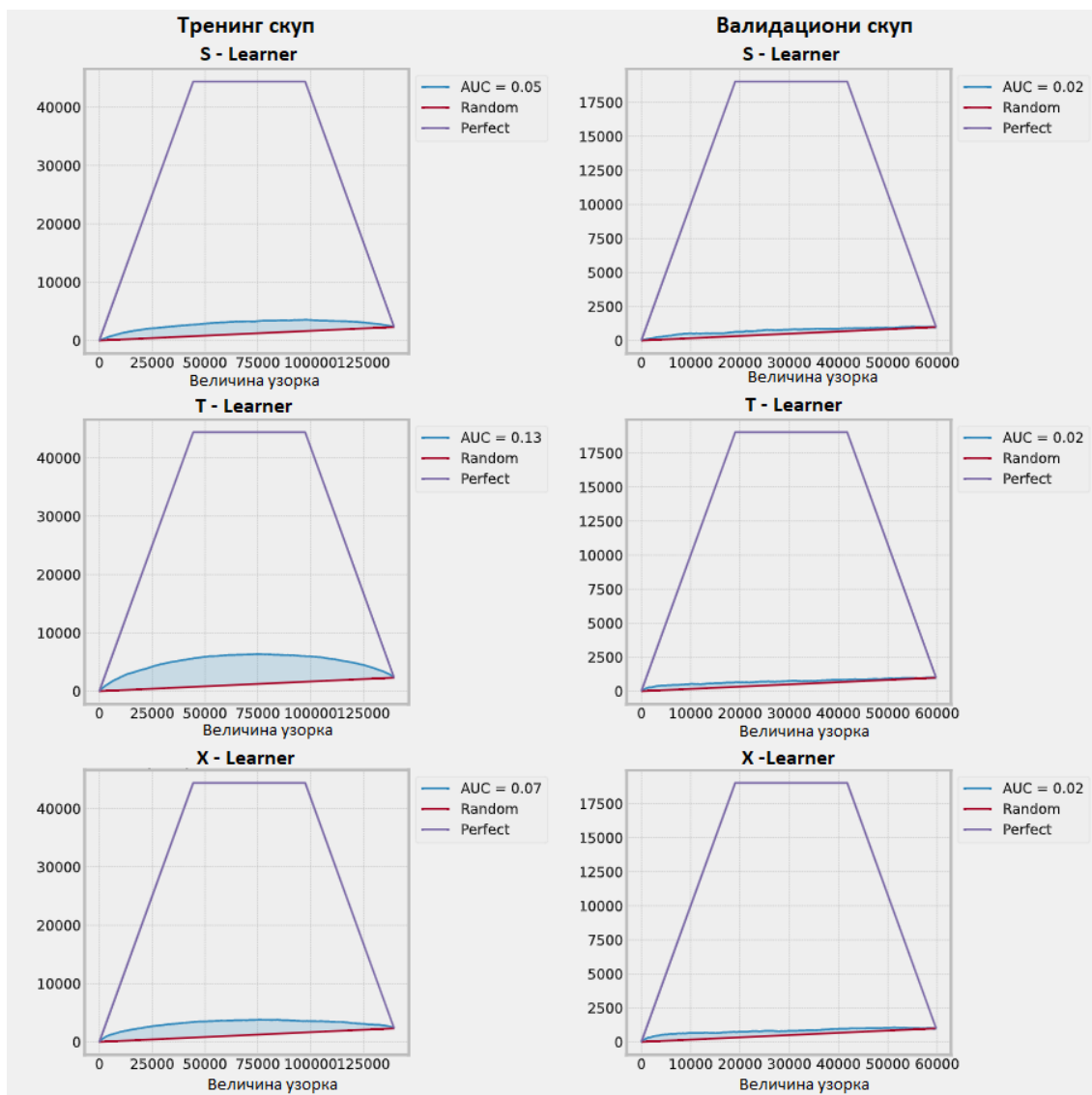
У конкретном примеру, чини се да је T-Learner модел имао најбоље перформансе на тренинг скупу, док је X-Learner имао за нијансу боље перформансе на тест скупу што је очекивано зато што овај модел на ефикаснији начин комбинује информације о контролној и третманској групи.

У раду са реалним подацима, тешко је постићи савршен ефекат уздизања, па је у том смислу мање смислено поредити криве уздизања са кривом савршеног модела. Бољу информацију о ефикасности модела нам даје поређење са кривом насумичног модела.

### **Кинијева крива**

Још једна описана мера евалуације модела је Кинијева крива, коју смо увели у једначини (4.3).

Кинијева крива је слична претходно описаној кривој уздизања, али пружа директније поређење кумулативног утицаја циљања клијената на основу предвиђања ефекта уздизања, у односу на насумично циљање. У пракси се користе обе криве које нам дају сличну информацију. Исти закључак о ефикасности модела се односи и на овај график.



Слика 6.7: Кинијева крива за сва 3 модела

## Дискусија

Циљ овог поглавља је био да прикажемо моделовање уздицања на реалним подацима. Као што видимо са графикана, сва 3 представљена модела нису толико добрих перформанси, али за потребе побољшања модела, постоје додатни алати и функције унутар описаног пакета и библиотеке.

За потребе одлучивања које клијенте бисмо контактирали, намеће се закључак да треба посматрати криву побољшања и наћи вредност у којој крива достиже максимум, и до те вредности контактирати клијенте, како не бисмо

трошили непотребна финансијска средства за кампању.

Такође на основу резултата модела који нам даје предвиђене вредности индивидуалног ефекта третмана, можемо клијенте да класификујемо у једну од три категорије које су нам од значаја: клијенти склони убеђивању, скептични клијенти и клијенти без утицаја. Ова информација нам је од великог значаја за све будуће кампање за које ће се користити овај скуп података јер ћемо унапред знати коју категорију клијената је најефективније контактирати.

## Глава 7

### Закључак

У овом раду је приказано више различитих приступа за оцењивање узрочних ефеката. Први је дводелни приступ који се састоји од обучавања два одвојена модела: једног за третирану групу, а други за контролну групу. Ефекат уздицања тест података се израчунава као разлика између његове предикције у два модела. Иако је дводелни приступ једноставан за примену, он може бити превазиђен приступом трансформације класе који има за циљ моделирање трансформисане циљне променљиве чије је условно очекивање једнако стварном ефекту уздицања. Други метод се ослања на претпоставку потпуне рандомизације третмана. Трећи приступ подразумева модификовање постојећих модела машинског учења. У овом раду смо ограничили пажњу на методе засноване на алгоритмима дрвећа и представили различите критеријуме за поделу.

У вези евалуације модела, запажено је да у одсуству тачног ефекта уздицања, нема лако израчунљивог губитка који би могао да се користи за процену перформанси модела. Један од приступа подразумева сортирање третираних и нетретираних тест података, по растућем редоследу предвиђеног ефекта уздицања, одвојено. Обе групе се потом разврставају у дециле, након чега се перформансе модела процењују кроз међусобну разлику у просечном ефекту уздицања по децилима. Прецизнија метода процене која је заправо генерализација међусобне оцене кумулативних разлика, јесте крива уздицања. Тестне опсервације се сортирају по растућем редоследу предвиђеног позитивног ефекта. Крива уздицања је дефинисана као параметарска функција броја одабраних опсервација која враћа разлику у просеку предвиђеног позитивног ефекта између третиране и контролне групе. Крива уздицања обично има

облик звона, а површина испод те криве служи као мера перформанси (као што је то случај са традиционалном AUC). Коначно, у овом раду смо представили методу која директно користи трансформисану циљну варијаблу  $Y_i^*$  и у једначини MSE. Доказали смо да, у граници, минимизирање формуле MSE која користи  $Y_i^*$  и уместо тачног позитивног ефекта такође минимизира једначину MSE која користи тачан позитивни ефекат.

Поред поменутих метода, приказали смо преглед алгоритама мета-модела из библиотеке *Causal ML* које смо применили на реалним подацима и приказали њихове резултате.

# Библиографија

- [1] Aja19. causal-inference-and-uplift-modeling. <https://github.com/Aja19/causal-inference-and-uplift-modeling/tree/main>.
- [2] Joshua D Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2009.
- [3] Eva Ascarza. Retention futility: Targeting high-risk customers might be ineffective. *Journal of Marketing Research*, 55(1):80–98, 2018.
- [4] Susan Athey and Guido Imbens. Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360, 2016.
- [5] Susan Athey and Guido W Imbens. Machine learning methods for estimating heterogeneous causal effects. *stat*, 1050(5):1–26, 2015.
- [6] Susan Athey and Guido W Imbens. The econometrics of randomized experiments. In *Handbook of economic field experiments*, volume 1, pages 73–140. Elsevier, 2017.
- [7] Leo Breiman. Random forests. *Machine learning*, 45:5–32, 2001.
- [8] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.
- [9] CodeEmporium. Causal inference and machine learning - explained. <https://www.youtube.com/watch?v=0d6oAz10p2k&list=PLGujnZcegunexcxy20SIZ1nYvg1p8m8M4&index=4>.
- [10] Eustache Diemert, Artem Betlei, Christophe Renaudin, and Massih-Reza Amini. A large scale benchmark for uplift modeling. In *KDD*, 2018.

- [11] Leo Guelman, Montserrat Guillén, and Ana M Pérez-Marín. Uplift random forests. *Cybernetics and Systems*, 46(3-4):230–248, 2015.
- [12] Leo Guelman, Montserrat Guillén, and Ana María Pérez Marín. Optimal personalized treatment rules for marketing interventions: A review of methods, a new proposal, and an insurance case study. *UB Riskcenter Working Paper Series, 2014/06*, 2014.
- [13] Pierre Gutierrez and Jean-Yves Gérardy. Causal inference and uplift modelling: A review of the literature. In *International conference on predictive applications and APIs*, pages 1–13. PMLR, 2017.
- [14] Behram Hansotia and Brad Rukstales. Incremental value modeling. *Journal of Interactive Marketing*, 16(3):35–46, 2002.
- [15] Szymon Jaroszewicz and Piotr Rzepakowski. Uplift modeling with survival data. In *ACM SIGKDD Workshop on Health Informatics (HI-KDD-14)*, New York City, 2014.
- [16] Maciej Jaskowski and Szymon Jaroszewicz. Uplift modeling for clinical trial data. In *ICML Workshop on Clinical Data Analysis*, volume 46, pages 79–95, 2012.
- [17] Finn C Kuusisto. *Machine learning for medical decision support and individualized treatment assignment*. PhD thesis, The University of Wisconsin-Madison, 2015.
- [18] Lily Yi-Ting Lai. Influential marketing: a new direct marketing strategy addressing the existence of voluntary buyers. 2006.
- [19] Victor SY Lo. The true lift model: a novel data mining approach to response modeling in database marketing. *ACM SIGKDD Explorations Newsletter*, 4(2):78–86, 2002.
- [20] Oscar Mesalles Naranjo. Testing a new metric for uplift models. 2012.
- [21] Houssam Nassif, Finn Kuusisto, Elizabeth S Burnside, and Jude W Shavlik. Uplift modeling with roc: An srl case study. In *ILP (late breaking papers)*, pages 40–45. Citeseer, 2013.

- [22] Nicholas Radcliffe. Using control groups to target on predicted lift: Building and assessing uplift model. *Direct Marketing Analytics Journal*, pages 14–21, 2007.
- [23] Nicholas J Radcliffe. Hillstrom’s minethatdata email analytics challenge: An approach using uplift modelling. *Stochastic Solutions Limited*, 1:1–19, 2008.
- [24] Nicholas J Radcliffe and Patrick D Surry. Real-world uplift modelling with significance-based uplift trees. *White Paper TR-2011-1, Stochastic Solutions*, pages 1–33, 2011.
- [25] Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.
- [26] Piotr Rzepakowski and Szymon Jaroszewicz. Decision trees for uplift modeling with single and multiple treatments. *Knowledge and Information Systems*, 32:303–327, 2012.
- [27] scozahulio. Uplift modeling. <https://www.kaggle.com/datasets/scozahulio/upliftm>.
- [28] Atef Shaar, Talel Abdesslem, and Olivier Segard. Pessimistic uplift modeling. *arXiv preprint arXiv:1603.09738*, 2016.
- [29] Michał Sołtys, Szymon Jaroszewicz, and Piotr Rzepakowski. Ensemble methods for uplift modeling. *Data mining and knowledge discovery*, 29:1531–1559, 2015.
- [30] Xiaogang Su, Joseph Kang, Juanjuan Fan, Richard A Levine, and Xin Yan. Facilitating score and causal inference trees for large observational studies. *Journal of Machine Learning Research*, 13:2955, 2012.
- [31] Stefan Wager and Susan Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 2018.
- [32] Lukasz Zaniewicz and Szymon Jaroszewicz. Support vector machines for uplift modeling. In *2013 IEEE 13th International Conference on Data Mining Workshops*, pages 131–138. IEEE, 2013.



# Биографија аутора

Александра Поповић је рођена 19. марта 1998. године у Београду. Завршила је основну школу „Душко Радовић” у Београду, а након тога Девету гимназију „Михајло Петровић Алас”. 2017. године уписује Математички факултет Универзитета у Београду, на смеру статистика, актуарска и финансијска математика. Основне студије завршава 2021. године са просечном оценом 9.03. Главне области интересовања су јој вероватноћа и статистика, као и машинско учење.