

# Detekcia anomálií v časových radoch: prevencia gamblingu pomocou hlbokého učenia

Bibiána Lajčinová<sup>1</sup>, Marián Gall<sup>1,2</sup> and Michal Pitoňák<sup>1,3</sup>

<sup>1</sup>Národné superpočítačové centrum, Dúbravská cesta 3484/9, 84104 Bratislava, Slovenská republika

<sup>2</sup>Ústav informatizácie, automatizácie a matematiky, Fakulta chemickej a potravinárskej technológie, Slovenská technická univerzita v Bratislave, Radlinského 9, 812 37 Bratislava, Slovenská republika

<sup>3</sup>Katedra fyzikálnej a teoretickej chémie, Prírodovedecká fakulta, Univerzita Komenského v Bratislave, Iľkovičova 6, 842 15 Bratislava, Slovenská republika

## Abstrakt

*Prevencia gamblingu u hráčov online kasín je výzvou, so zjavne pozitívnym dopadom nielen na bežný život hráčov, ale aj na prevádzkovateľov kasín, ktorých úmyslom je sprostredkovať zodpovedné hranie. Na dosiahnutie týchto cieľov sme použili metódy hlbokého učenia bez učiteľa (z angl. "unsupervised learning"), ktoré dokážu identifikovať hráčov vykazujúcich známky problémového hráčstva, s využitím dostupných dát vo forme časových radov. Prezентujeme porovnanie nami navrhnutej architektúry autoenkódera založenej na transformeroch, s rekurentnými a konvolučnými autoenkódermi, s dôrazom na výhody tejto architektúry v detekcii anomálií. Keďže klinická diagnóza hráčov nebola v dostupných dátach k dispozícii, výsledky našej štúdie vyhodnocujeme analyzovaním "skóre anomálie" získaného z autoenkódera a niekoľkých pomocných ukazovateľov, ktoré sú v literatúre často spomínané ako symptómy problémového hráčstva.*

## Úvod

Patologické hranie hier je definované ako problémové správanie, kedy osoba opakovane stavia objekt určitej hodnoty (najčastejšie peniaze) za účelom výhry, ale s neurčitým výsledkom [1], [2]. Bolo tiež odhalených mnoho podobností medzi problémovým gamblingom a závislosťami na omamných látkach [3]. S prístupnosťou internetu sa zvýšil aj výskyt patologického hráčstva. Toto správanie má často negatívny dopad nielen na život dotknutých jednotlivcov, ale aj ich rodiny. Práve preto je včasná detekcia známk patologického hráčstva kľúčová nielen pre zachovanie duševného zdravia, ale aj finančnej situácie ovplyvnených osôb.

Táto práca vznikla ako spolupráca Národného kompetenčného centra pre vysokovýkonné počítanie, spoločnosti DOXXbet s.r.o. (športové tipovanie a online kasíno) a spoločnosti Codium s.r.o. (vývoj softvérovej platformy športové tipovanie a iGaming-u pre DOXXbet, s.r.o.), s cieľom zlepšenia služieb zákazníkom a prevencie patologického hráčstva.

Tento "proof of concept" projekt je základom pre ďalší vývoj nástrojov, ktoré pomôžu poskytovateľovi zmierniť dopad negatívnych

následkov problémového hráčstva na životy hráčov, aj za cenu zníženia príjmov, čo je v súlade s aktuálnymi európskymi trendmi v risk manažmente online hazardných hier.

Existuje viacero štúdií zaoberajúcich sa detekciou hráčov, ktorí prejavujú problémy so zvládaním hazardných hier s použitím metód strojového učenia. Väčšina týchto štúdií však využíva znalosť klinickej diagnózy jednotlivcov, vďaka čomu je možné použiť metódy učenia s učiteľom (z angl. "supervised learning"). Bez tejto informácie a taktiež bez informácie o finančných problémoch hráča sa musíme spoliehať len na pomocné ukazovatele, medzi ktorými sú napríklad dobrovoľné vylúčenie z online kasína. Medzi ďalšie vedľajšie príznaky spomínané v literatúre patria sťažnosti sprostredkovateľovi online kasína, žiadosti o zvýšenie limitu [4], doháňanie strát intenzívnejším hraním, alebo napríklad preferovanie okamžitej odmeny (z angl. "instant gratification", t.j. správanie, kedy jednotlivec preferuje nižšiu odmenu okamžite, miesto vyššej odmeny s odstupom času) a mnoho ďalších [5], [6]. S použitím iba spomenutých údajov o správaní hráča nie je, samozrejme, možné s

istotou povedať, či má hráč problém s hazardnými hrami. Napriek tomu sa však v tejto štúdií pokúsime o detekciu skorých príznakov patologického hráčstva, ktoré môžu byť užitočné aj za cenu falošne pozitívnych predpovedí.

Jedna zo súčasných štúdií, z ktorej táto práca výrazne čerpá, sa zameriava na identifikáciu hráčov s potenciálnym patologickým hráčstvom, pričom sa zameriava na hry založené na "zručnostiach" (nie na náhode) [7] a používa na to neuronové siete s dlhodobým a krátkodobým pamäťovým efektom (LSTM, z angl. "Long Short-Term Memory") a adverzálne autoenkodery (AAE). Autori tohto článku disponovali klinickou diagnózou hráčov, vďaka čomu neboli nútení spoliehať sa na pomocné ukazovatele. Iná štúdia [8] využíva metódu segmentovania používateľov online kasína podľa sklonov ku kompulzívnemu hraniu a pracuje, rovnako ako naša štúdia, s neklasifikovanými dátami.

V našej práci sme použili prístup učenia bez učiteľa s využitím neuronovej AE siete založenej na transformeroch za účelom detekcie anomálií. Predpokladáme, že dáta patologických hráčov, ktorých by (štatisticky) medzi všetkými hráčmi malo byť medzi 0.5 až 5% [7], vykazujú anomálie, nie nutne naopak. Dátová sada, ktorú používame, neobsahuje klinickú diagnózu hráčov a spomedzi pomocných ukazovateľov spomenutých vyššie sa v dátach nachádzajú len niektoré – žiadosti o zvýšenie limitu vkladov, doháňanie strát intenzívnejším hraním (ďalej budeme tento indikátor označovať "epizódy naháňania straty"), používanie viacerých platobných metód, časté výbery malých finančných obnosov a ďalšie, ktoré budú spomenuté neskôr. Nakoľko nie všetci hráči, ktorí preukážu anomálne správanie musia byť patologickými hráčmi, boli pomocné ukazovatele použité v kombinácii so *skóre anomálie*, získaného ako výstup z AE.

## Dáta

Dáta použité v tejto štúdií sú sekvenciami pozorovaní zbieraných v čase, monitorujúce viaceré aspekty hráčskeho správania ako frekvencia a časy hier, frekvencia a výška peňažných vkladov, platobné metódy použité pri peňažných vkladoch, informácie o výške stávok, výhier, strát,

výberoch a taktiež početnosť žiadostí o zmenu limitu pre výšku vkladov. Z týchto dát bolo extrahovaných 19 prvkov deskriptora vo forme časových radov (ČR). Každý prvok deskriptora pozostáva z viacerých hodnôt, ktoré sú chronologicky usporiadané. Tieto príznaky sú pre zrozumiteľnosť klasifikované do troch kategórií – "čas", "peniaze" a "zúfalstvo", inšpirované autormi Seth *et al.* [7]. Tabuľka 1 sumarizuje kompletnú sadu použitých premenných s krátkym popisom.

Každý prvok deskriptora je sekvenciou  $N$  hodnôt, pričom každá hodnota predstavuje jedno z  $N$  po sebe idúcich časových okien. Každá takáto hodnota bola získaná agregáciou denných dát v príslušnom časovom okne, pričom dĺžka časového okna je špecifikovaná v Tabuľke 1, spolu s informáciou, či ide o pohyblivé časové okno alebo nie. Pre každé pozorovanie potrebujeme dáta z  $N$  časových okien. Celý proces extrakcie premenných je zobrazený na Obrázku 1, finálny tvar dát na Obrázku 2.

Dĺžka sekvencie  $N$  hrá dôležitú úlohu v kvalite predikcie. Zatiaľ čo vyššia hodnota  $N$  prináša vo všeobecnosti lepšie výsledky, zvyšovanie tejto hodnoty je v rozpore s našim cieľom detegovať anomálie v čo najkratšom možnom čase. Preto je preferovaná čo najnižšia hodnota  $N$ . Po sérii experimentov bola zvolená hodnota  $N = 8$ . Na získanie premenných boli použité tri agregčné funkcie: maximum, minimum a/alebo suma (bližšie špecifikované v Tabuľke 1).

Pre niektoré premenné z Tabuľky 1 uvádzame detailnejší popis:

- *Money 10* – Počet dní v časovom okne, kedy hráč urobil aspoň dva malé výbery za jeden deň. Ide o dni, kedy si hráč vybral peniaze aspoň dvakrát za deň, pričom suma hodnôt týchto výberov bola menšia, než tridsiaty percentil z hodnôt súm výberov hráčov, ktorí vykonali aspoň dva výbery denne.
- *Chasing ep. 1 & 2* – Počet dní v časovom okne, kedy hráč naháňal stratu. Epizóda naháňania straty je udalosť, ktorá nastane, keď má hráč lokálne maximum strát sprevádzané lokálnym maximom vkladov v časovom okne a tento extrém peňažných vkladov pozostáva aspoň z troch úkonov.

Kontext	Názov	Časový rámec	Popis
Čas	Time 1	15-dňové PO	Max. počet hier odohraný za deň v časovom okne.
	Time 2	15-dňové PO	Celkový počet hier hraných cez víkend v časovom okne.
	Time 3	15-dňové PO	Celkový počet hier hraných v prac. deň v časovom okne.
	Time 4	15-dňové PO	Celkový počet hier hraných medzi polnocou a 06:00 v časovom okne.
	Time 5	15-dňové PO	Celkový počet prihlásení v časovom okne.
Peniaze	Money 1	1-mes. PO	Celková strata, kt. vznikla za posledný mesiac.
	Money 2	3-mes. PO	Celková strata za posledné 3 mesiace.
	Money 3	15-dňové PO	Celková strata za časové okno.
	Money 4	15-dňové PO	Maximálny počet denných vkladov za časové okno.
	Money 5	15-dňové PO	Počet viacnásobných (> 3) denných vkladov v časovom okne.
	Money 6	15-dňové PO	Max. suma vložená za jediný deň v časovom okne.
	Money 7	15-dňové PO	Max. počet rôznych platobných metód použitých za deň v časovom okne.
	Money 8	15-dňové PO	Počet dní v časovom okne, kedy boli použité viac než dve platobné metódy.
	Money 9	15-dňové PO	Počet výberov v časovom okne.
	Money 10	15-dňové PO	Počet dní v časovom okne, kedy hráč urobil aspoň dva malé výbery za jeden deň.
Zúfalstvo	Despair 1	15-dňové PO	Min. pomer výhier k celkovému počtu hier za deň v časovom okne.
	Despair 2	15-dňové PO	Počet žiadostí o zvýšenie denného limitu vkladov v časovom okne.
	Chasing ep. 1	15-dňové PO	Počet dní v časovom okne, kedy hráč "naháňal stratu".
	Chasing ep. 2	15-dňové ČO	Počet dní v časovom okne, kedy hráč "naháňal stratu".

**Tabuľka 1:** Zoznam odvodených prvkov deskriptora vo forme časových radov. Skratky s stĺpci "Časový rámec": PO - pohyblivé (časové) okno, ČO - disjunktné časové okno.

Na Obrázku 3 je uvedený príklad vstupných dát pre náhodne vybraného hráča z dátovej sady, pozostávajúci z 19 prvkov deskriptora, t.j. 19 časových radov.

## Čistenie dát

Identifikácia a odstraňovanie abnormálnych vstupných dát z trérovacej sady pri použití AE na detekciu anomálií je dôležitým krokom na zabezpečenie toho, aby sa model učil hlavne na dátach, ktoré sú "normálne" [9]. Dáta, ktoré používame, takmer určite obsahujú anomálie, keďže sa jedná o dátovú sadu zahrňujúcu všetkých hráčov bez filtrovania, takže aj suspektných gamblerov.

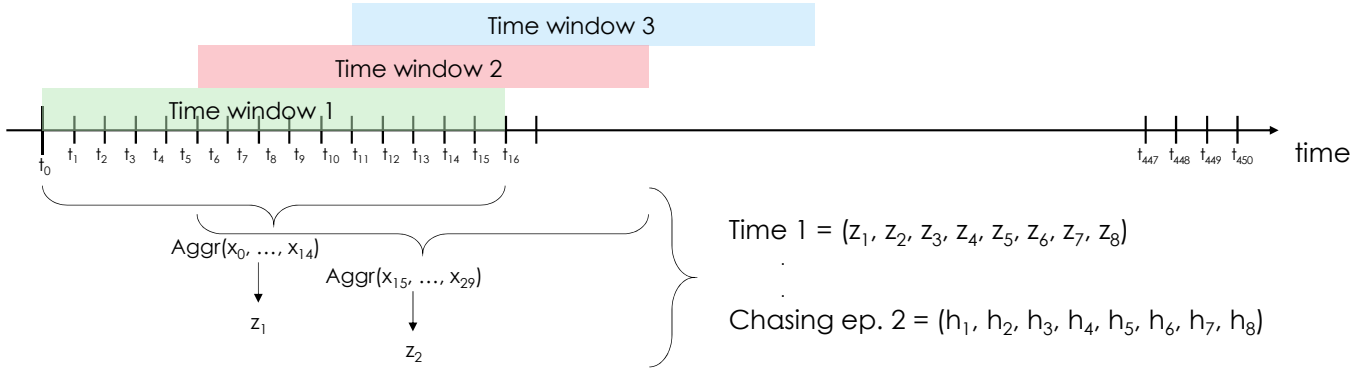
Náš prístup k čisteniu dát pozostáva z použitia pevne stanovenej sady pravidiel. Tieto pravidlá môžu samy o sebe byť principiálne použité aj ako jednoduchá metóda na detegovanie abnormálnych hráčov, nedosahujú však zd'aleka presnosť AE modelov. Pravidlá sú založené už na spomenutých pomocných ukazovateľoch, konkrétne nasledovných: počet žiadostí o zmenu limitu výšky vkladov, priemerný počet prihlásení hráča, priemerný počet výberov hráča a

priemerný počet epizód naháňania straty. Ako hranicu pre rozhodovanie, či dané pozorovanie je alebo nie je anomália, použijeme 95-ty percentil z hodnôt príslušných pomocných ukazovateľov. Ak hráč splňa aspoň tri z týchto štyroch pomocných ukazovateľov (t.j. hodnoty sú vyššie ako stanovené hranice), hráča považujeme za anomálne pozorovanie a z trérovacej sady je presunutý do testovacej. Spolu bolo z trérovacej sady odstránených 211 pozorovaní, čo predstavuje približne 1% z jej veľkosti.

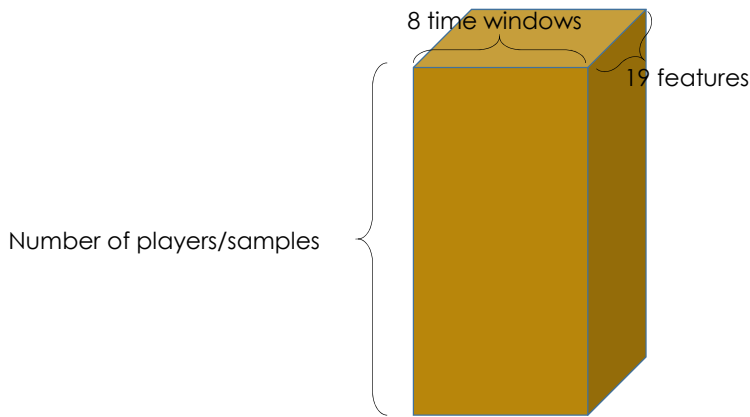
## Normalizácia dát

Po spracovaní "surových dát" a ich následnom čistení, nasleduje normalizácia, ktorá je dôležitá kvôli optimálnemu priebehu trérovania modelu. Normalizácia je proces úpravy hodnôt v dátovom súbore tak, aby mali všetky rovnaký rozsah. To môže pomôcť zlepšiť presnosť modelu a takisto optimalizuje proces učenia. *Z-Score normalizácia* bola vybraná spomedzi normalizačných techník po sérii experimentov s ďalšími technikami ako *log* a *minmax* normalizácia.

**Z-Score normalizácia** prebieha nasledovne: Nech  $\mathbf{x} \in \mathbb{R}^D$  je vektor  $\mathbf{x} = (x_1, \dots, x_D)$ . Najskôr



**Obr. 1:** Vizualizácia agregácie dát na dennej báze do formy časových okien a finálnej formy prvkov deskriptora.  $t_1, \dots, t_{450}$  predstavujú časové záznamy pre denné dáta  $x_1, \dots, x_{450}$ . Denné záznamy z časového okna sú agregované do jednej hodnoty  $z_i$  pre všetky  $i \in (1, \dots, 8)$ .



**Obr. 2:** Finálny tvar dát vstupujúcich do modelu. Každé pozorovanie pozostáva z devätnástich prvkov deskriptora a každý prvok pozostáva z ôsmich časových okien.

vypočítame priemer a smerodajnú odchýlku  $x$ :

$$\mu_x = \frac{1}{D} \sum_{i=1}^D x_i \quad (1)$$

$$\sigma_x = \sqrt{\frac{1}{D} \sum_{i=1}^D (x_i - \mu_x)^2} \quad (2)$$

Z-score normalizácia  $x$  je potom vypočítaná ako:

$$ZN(x) = \frac{x - \mu_x \mathbf{1}}{\sigma_x} \in \mathbb{R}^D \quad (3)$$

kde  $\mathbf{1} = [1, \dots, 1]^T$  je  $D$ -dimenzionálny vektor jednotiek. Táto normalizácia má užitočnú vlastnosť, že transformuje odľahlé pozorovania na štandardizované hodnoty, ktoré sú porovnateľné so zvyškom dát, čím zjednodušuje analýzu a interpretáciu týchto dát.

Pre premenné, ktoré vykazovali vysoký rozptyl bola normalizácia zrealizovaná na úrovni pozorovaní/hráčov, zatiaľ čo pre ostatné premenné, ktoré vykazovali nízky rozptyl bola normalizácia urobená na úrovni celej dátovej množiny. Premenné, ktoré neboli normalizované na úrovni hráča sú *despair 1* a *despair 2*.

### Data clipping

Tzv. "data clipping" je ďalšia technika použitá na spracovanie dát spolu so Z-Score normalizáciou. Nakoľko pri čistení dát bolo odstránených iba 211 pozorovaní, je potrebné eliminovať vplyv zvyšných odľahlých pozorovaní pomocou data clippingu. Tento krok bol vykonaný s použitým metriky nazývanej MAD (z angl. "Median Absolute Deviation") definovanej v rovnici 4.

$$MAD = \text{median}(|x_i - \tilde{x}|) \quad (4)$$

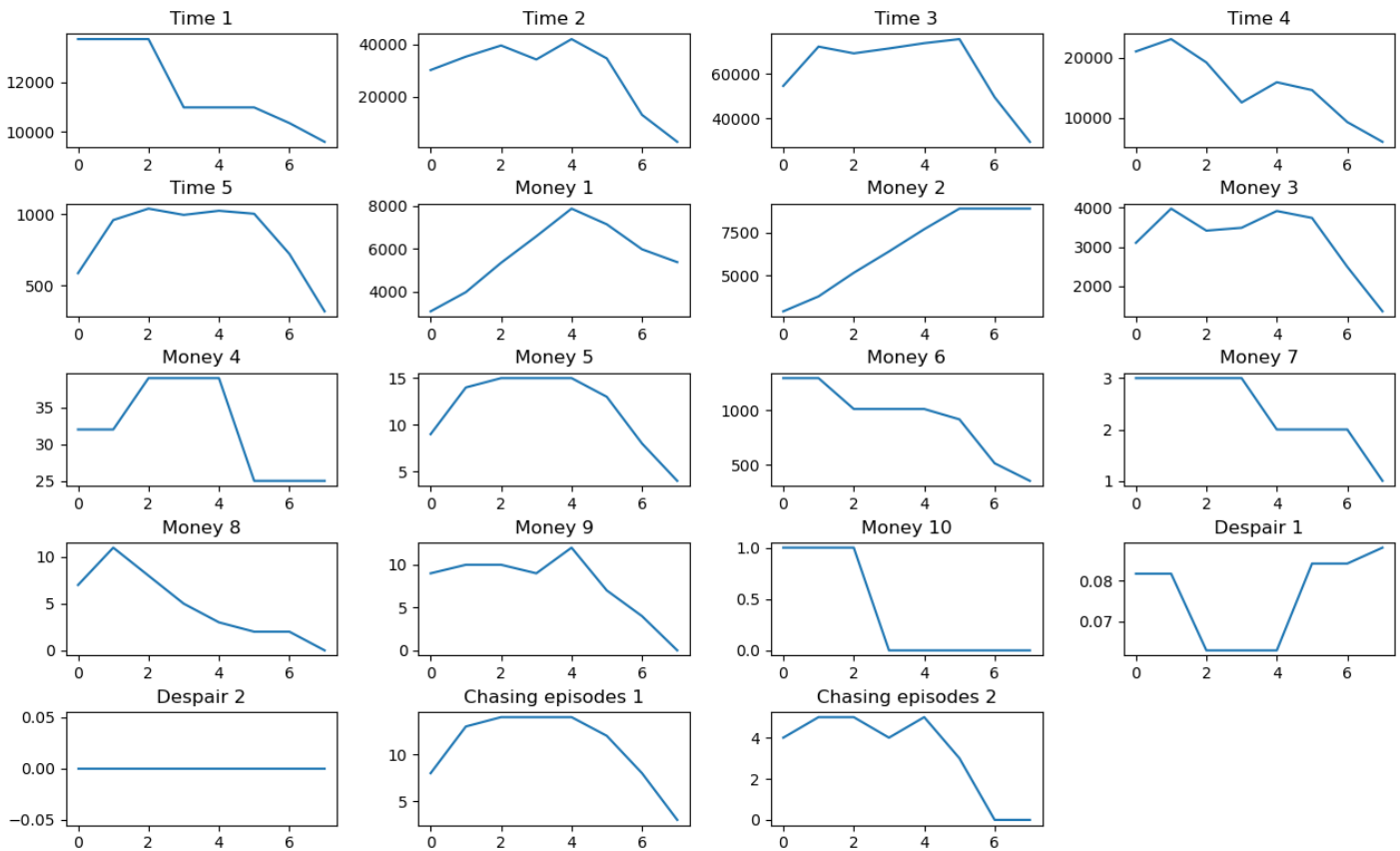
kde  $\tilde{x}$  reprezentuje medián hodnôt  $x_1, x_2, \dots, x_d$ . MAD meria priemernú vzdialenosť medzi každým pozorovaním a mediánom celej dátovej sady. Ak pre nejaké pozorovanie platí

$$|x_i| > K * MAD \quad (5)$$

kde  $K$  je konštanta, tak hodnota  $x_i$  je považovaná za odľahlé pozorovanie a je "ostrihnuté" nasledovným spôsobom

$$x_i = x_i - MAD \quad (6)$$

Vybraná hodnota pre  $K$  bola 4.



**Obr. 3:** Príklad vstupných dát. Každý prvok deskriptora pozostáva z 19 časových radov, pričom každý časový rad predstavuje jednu premennú. Každá premenná pozostáva z ôsmich časových okien a každé časové okno reprezentuje 15 dní. Zobrazené hodnoty nie sú škálované, ani neprešli procedúrou "data clipping-u", ktoré sú nevyhnutné pred vstupom do modelu.

## Architektúra modelu neurónovej siete

Časový rad tvoria hodnoty, ktoré sledujú vývoj danej veličiny v čase, t.j. sú chronologicky usporiadané. Dáta tohto typu sa využívajú na predikciu budúcich hodnôt, ale taktiež na detekciu anomálií. Hlboké učenie nie je jediný prístup aplikovateľný na ČR, existujú aj klasické metódy ako napr. STL dekompozícia (z angl. "Seasonal-Trend Decomposition") [10], alebo regresné stromy.

AE je metóda hlbokého učenia bez učiteľa vhodná na detekciu anomálií v ČR. Myšlienka použitia tohto typu neurónovej siete na detekciu anomálií je založená na rekonštrukčnej schopnosti modelu. AE sa naučí rekonštruovať dáta v trénovacej sade a keďže trénovacia sada by ideálne mala obsahovať len "normálne" pozorovania, model sa naučí správne rekonštruovať

len takéto pozorovania. Preto keď je vstupné pozorovanie anomálne, natrénovaný AE model nedokáže tento vstup zrekonštruovať dostatočne správne, čo má za následok vysokú rekonštrukčnú chybu. Táto rekonštrukčná chyba môže byť použitá ako skóre anomálie daného pozorovania, pričom vyššie skóre znamená väčšiu pravdepodobnosť, že pozorovanie sa vymyká všeobecnému trendu.

AE môžeme ďalej rozdeliť do troch kategórií podľa toho, aký typ vrstiev/architektúru využívajú [11].

Prvou kategóriou sú LSTM AE. LSTM je typ rekurentnej neurónovej siete, ktorá bola navrhnutá práve pre sekvenčné dáta. Ďalší, bežne používaný typ neurónovej siete je CNN (Convolutional Neural Network). Na spracovanie časových radov je možné použiť 1D konvolučné vrstvy, alebo transformovať časový rad do 2- alebo 3D reprezentácie a následne použiť viacrozmerné konvolučné vrstvy.

Poslednou je architektúra neurónovej siete AE založená na transformeroch. Ide o pomerne novú triedu modelov hlbokého učenia, ktorá bola pôvodne vyvinutá pre spracovanie prirodzeného jazyka (NLP, z angl. "Natural Language Processing"), ale bolo preukázané, že poskytuje kvalitné výsledky aj pre iné typy sekvenčných dát [12], vrátane ČR.

LSTM a CNN neurónové siete sú často používané na prácu s ČR, avšak obe majú nedostatky v porovnaní s architektúrami založenými na transformeroch. LSTM a CNN architektúry neboli navrhnuté tak, aby zachytávali dlhodobé závislosti v dátach, čo môže limitovať ich prediktívnu schopnosť. Taktiež, tieto dve architektúry boli primárne navrhnuté na spracovanie sekvencií, a preto nedokážu dobre zachytiť nesekvenčné vzťahy v dátach (t.j. závislosti medzi dvoma bodmi, ktoré neležia vedľa seba / chronologicky za sebou), na rozdiel od transformerov, ktoré sú vhodné na modelovanie ako sekvenčných dát, tak aj nesekvenčných [13].

Pre každý typ AE architektúry existuje aj tzv. adverzálna verzia. Adverzálny AE (AAE) môže urobiť z AE generatívny model. V tomto prístupe sa AE trénuje voči dvom kritériám – prvým je "tradičné" kritérium rekonštrukčnej chyby a druhým je adverzálna tréningová kritérium [14]. AAE je taktiež model vhodný a používaný na detekciu anomálií.

## Náš prístup

Použitie AAE analogickým spôsobom ako v [7] sa ukázalo ako nevhodné pre dáta, ktoré sme mali k dispozícii. Dôvodom je, že AAE (rovnako ako generatívne adverzálna sieť) vyžadujú na spoľahlivú detekciu anomálií použiť generatívny AE a diskriminatívny diskriminátor, ktoré musia byť trénované striktne iba na pozorovaniach, ktoré nie sú anomálne [15], [16]. V našej dátovej množine, ako už bolo spomenuté, môžu byť (a takmer určite sú) zastúpené anomálne pozorovania.

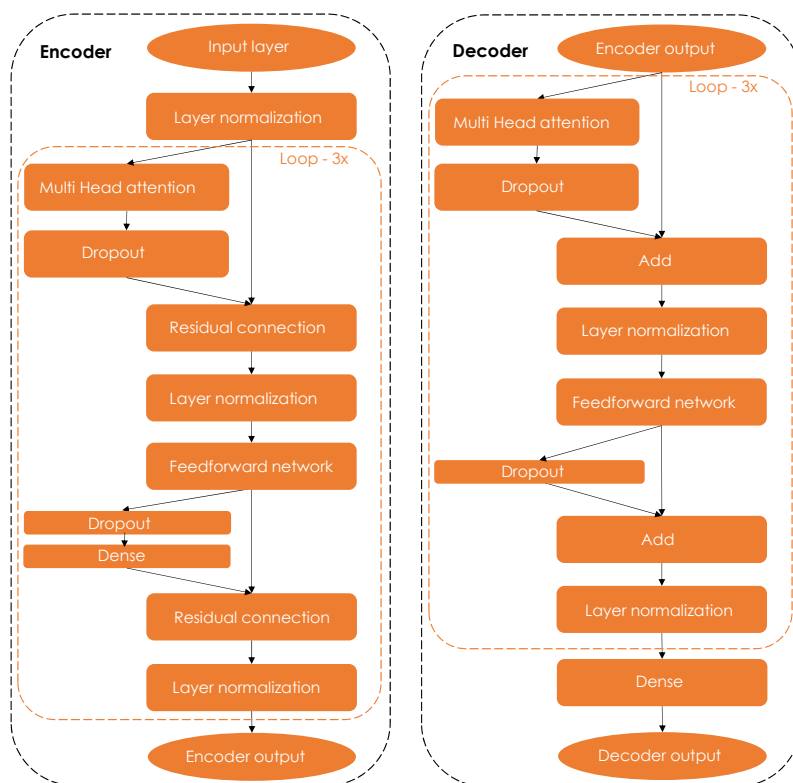
Aby sme mohli porovnať výsledky našej štúdie s výsledkami v [7], budeme používať architektúru AE navrhnutého v tomto článku, ale bez diskriminátora. Tento model budeme v ďalšom texte nazývať *LSTM A*. Ďalšie dve architektúry AE,

Názov modelu	Počet parametrov
<i>LSTM A</i>	2 435
<i>LSTM B</i>	691 163
<i>CNN</i>	291 299
<i>Transformer</i>	100 951

**Tabuľka 2:** Počet trénovateľných parametrov pre štyri porovnávané AE modely.

ktoré budeme testovať, sú: architektúra zložená z LSTM vrstiev popísaná v článku [17] (*LSTM B* model) a architektúra využívajúca 1D konvolučné vrstvy (*CNN* model). Posledný testovaný AE model je založený na transformeroch a budeme ho označovať *Transformer*. Tento model pozostáva z enkódera aj dekódera obsahujúceho vrstvu nazývanú *Multi Head Attention* so štyrmi "hlavami" a 32-dimenzionálnymi vektormi kľúčov  $K$  a hodnôt  $V$ . Za touto vrstvou nasleduje klasická neurónová sieť s *dropout* vrstvami a reziduálnymi spojeniami. Celý AE model má niečo málo viac než 100k trénovateľných parametrov. Vykonali sme niekoľko experimentov so zvyšovaním počtu parametrov modelu, ale ukázalo sa, že takáto hodnota je postačujúca, takže výsledky týchto experimentov v práci prezentované nie sú. Architektúra *Transformer* modelu je zobrazená na Obrázku 4. Jedným z hlavných zdrojov inšpirácie pre tento model bola ukážka kódu na Keras blogu [18]. Tento príspevok predstavuje architektúru modelu pre klasifikáciu časových radov s použitím transformerov a získal si našu pozornosť práve kvôli tomu, ako efektívne využíva "attention" mechanizmus.

Architektúry *LSTM A* a *LSTM B* modelov sú popísané v príslušných referenciách, kým architektúra *CNN* modelu bola navrhnutá nami. Enkóder aj dekóder obsahujú päť 1D konvolučných vrstiev, za ktorými nasledujú *batch normalization*, *dropout* a *maxpooling* vrstvy, s 256, 128, 64, 32 a 8 filtermi v konvolučných vrstvách enkódera a s rovnakým počtom filtrov, ale v opačnom poradí, vo vrstvách dekódera. Finálny počet parametrov každého zo štyroch AE modelov je sumarizovaný v Tabuľke 2.



Obr. 4: Vizualizácia architektúry neurónovej siete AE založeného na transformeroch.

## Porovnanie AE modelov

Všetky modely boli trénované na rovnakej dátovej množine s rovnakým nastavením (Adam optimalizátor, stratová funkcia – stredná kvadratická odchýlka (MSE)). Dátová sada obsahuje približne 22k pozorovaní (pričom každé pozorovanie pozostáva z 19 prvkov deskriptora vo forme ČR). Niektorí hráči sa nachádzajú v dátovej sade viackrát, pretože ich hráčska história je dostatočne dlhá na to, aby bolo možné vytvoriť viacero ČR s ôsmimi časovými oknami. V takomto prípade sú viaceré pozorovania prislúchajúce jednému hráčovi považované za nezávislé.

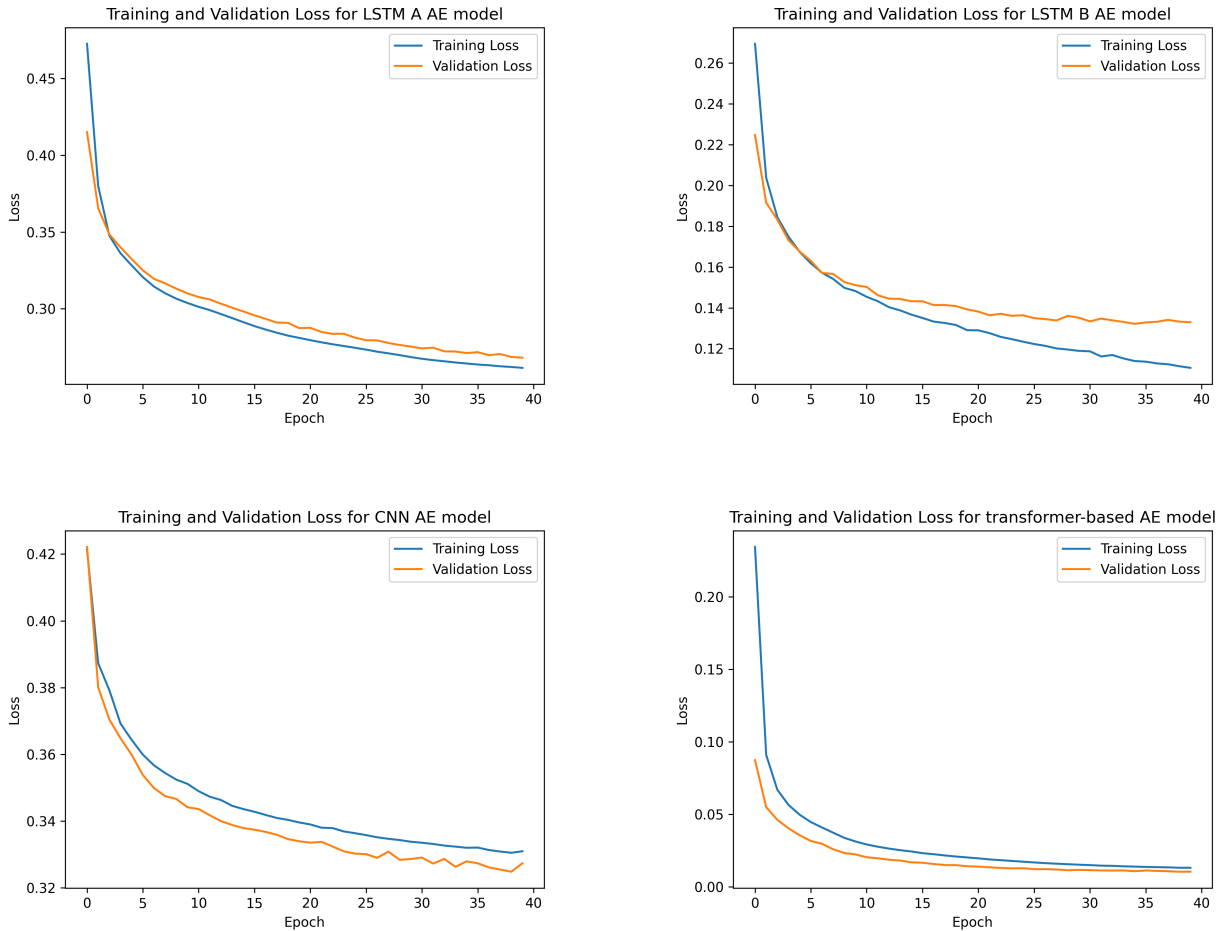
10% náhodne vybraných pozorovaní bolo použitých na validáciu a ďalších 10% na testovanie modelov. Dáta boli predspracované identicky pre každý zo štyroch trénovaných modelov (vrátane čistenia dát, "clippingu" a z-score normalizácie).

**Trénovacie krivky** História tréningu pre všetky štyri modely je zobrazená na Obrázku 5. Modré čiary zobrazujú hodnotu stratovej funkcie modelu na trénovacej množine a oranžové čiary na validačnej množine. Na osi x je počet

epoch použitých na tréning a na osi y je hodnota MSE straty.

Pre všetky modely strata rapídne klesá na začiatku tréningu. *Transformer* model konverguje najrýchlejšie zo všetkých štyroch modelov, približujúc sa stratou k nule bez preučenia (z angl. overfitting). *LSTM A* a *CNN* modely nevykazujú znaky pretrénovania, jedine *LSTM B* model prejavuje takéto sklony okolo desiatej epochy. Samotné straty dosiahnuté na konci tréningu sú najnižšie opäť pre *Transformer*, nasledované *LSTM B* modelom.

**Rekonštrukčná chyba** Na vyhodnotenie stability modelov sme uskutočnili 3-násobnú krížovú validáciu. Výsledné hodnoty priemernej straty, spolu s rozptylmi sú zobrazené v Tabuľke 3. Priemerná rekonštrukčná chyba modelu *Transformer* je značne nižšia než chyba ostatných modelov. *LSTM B* model nasleduje ako druhý v schopnosti rekonštruovať vstupné dáta, kým *CNN* model rekonštruuje najhoršie. Strata na testovacej množine je vždy vyššia než tá na trénovacej alebo validačnej množine, čo je spôsobené tým, že 211 anomálnych pozorovaní bolo presunutých do testovacej množiny. Bez presunu týchto pozorovaní by bola strata na testovacej množine pre



**Obr. 5:** Trénovalie krivky pre štyri AE modely. Na osi x je uvedený počet epôch a na osi y je hodnota straty meranej pomocou MSE.

Transformer model 0.012, pre CNN model 0.33, LSTMA model 0.27 a pre LSTMB model 0.13.

**Tabuľka 3:** Hodnoty stratovej funkcie pre štyri AE modely na trénovalnej (TR), validačnej (V) a testovacej (TS) množine získané z krížovej validácie. V prvom riadku sú uvedené hodnoty priemernej strednej kvadratickej odchýlky a v druhom riadku je rozptyl týchto odchýlok, pre každý z testovaných modelov.

Model	TR	V	TS
LSTM A	0.258	0.269	0.279
Rozptyl	4.2E-05	1.1E-05	1.4E-05
LSTM B	0.104	0.126	0.137
Rozptyl	1.4E-06	3.6E-06	1.2E-05
CNN	0.316	0.325	0.343
Rozptyl	9.8E-05	1.1E-05	1.1E-04
Transformer	0.010	0.012	0.015
Rozptyl	3.5E-06	1.4E-06	5.0E-06

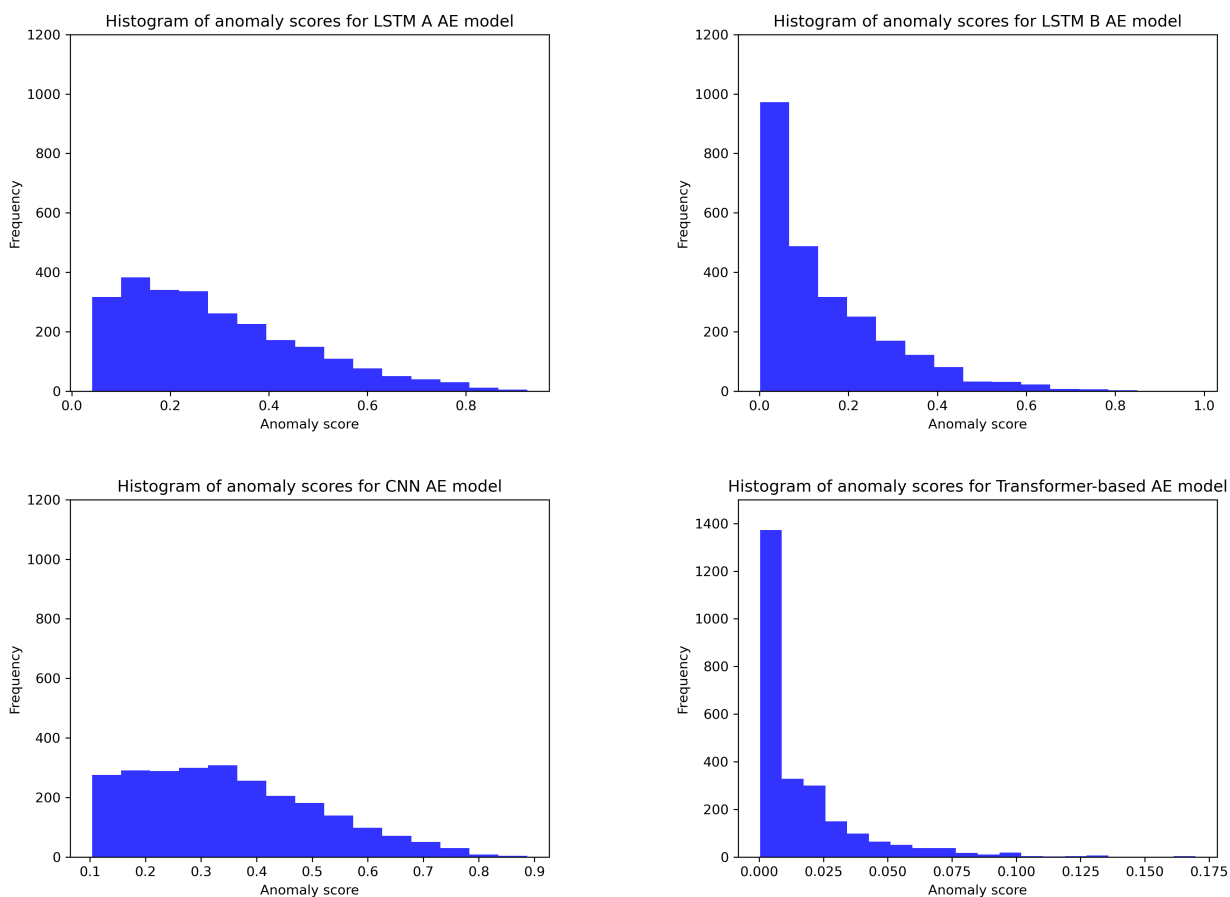
Detailnejší prehľad porovnania rekonštrukčnej schopnosti modelov je na Obrázku 6, pomô-

cou histogramu MSE hodnôt. Všetky histogramy majú "tážký pravý chvost", čo je očakávané pre dátové sady obsahujúce anomálie. Podobne ako na Obrázku 5, aj tu je zrejmé, že Transformer má výrazne nižšie hodnoty rekonštrukčnej chyby (os x) v porovnaní s ostatnými modelmi.

**Predikčná schopnosť** Pre demonštračné účely schopnosti rekonštrukcie ČR, Obrázok 8 zobrazuje pôvodné (modrá čiara) a predikované (červená čiara) hodnoty pre náhodne vybrané anomálne pozorovanie jedného hráča. Hodnota skóre anomálie pre príslušné modely je uvedená v nadpise grafov.

Rekonštrukčná schopnosť Transformeru je výrazne lepšia v porovnaní s ostatnými modelmi. Pripomíname, že tento model nie je preučovaný a toto pozorovanie pochádza z testovacej množiny. CNN a LSTM AE modely pre dáta vo forme ČR potrebujú podstatne dlhšiu sekvenciu dát





**Obr. 6:** Histogramy rekonštrukčnej chyby pre testovaciu množinu pre štyri porovnávané AE modely. Na osi x je hodnota skóre anomálie a na osi y je frekvencia príslušnej hodnoty.

/ historický kontext, aby dokázali dostatočne rekonštruovať vstupné ČR. *LSTM B* AE rekonštruuje výrazne lepšie než *LSTM A* model, ale stále nie tak kvalitne ako *Transformer*.

**Detegované anomálie** Posledné porovnanie modelov predstavuje kontrolu zhody medzi detegovanými množinami anomálií. Konkrétne, v testovacej množine boli identifikované pozorovania s najvyšším skóre anomálie (MSE) tak, že boli vybrané pozorovania, ktorých skóre anomálie patrí medzi posledných päť percentilov tejto hodnoty pre každý z modelov. Následne bola vyhodnotená zhoda v týchto identifikovaných množinách. Výsledky sú zobrazené v Tabuľke 4. Najväčšia zhoda je medzi *CNN* a *LSTM A* modelom, kde sa zhoduje takmer 70% anomálií identifikovaných týmito modelmi.

Prienik pozorovaní detegovaných všetkými štyrmi modelmi je zobrazený na Obrázku 7. Množina anomálií identifikovaná každým AE

modelom obsahuje 126 pozorovaní (hráčov) a v ich prieniku sa nachádza 34 hráčov.

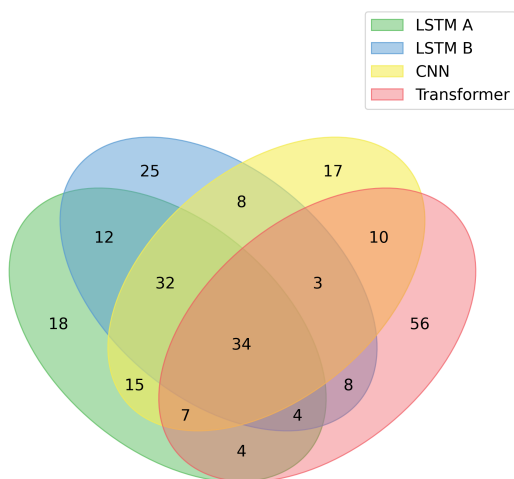
Vo zvyšku článku budeme diskutovať už len nami navrhnutý *Transformer AE* model, nakoľko bola preukázaná jeho mimoriadna schopnosť rekonštruovať dáta a preto očakávame, že tomu bude odpovedať aj jeho schopnosť detekcie anomálií.

## Výsledky modelu Transformer

Ako už bolo diskutované v úvode, klinická diagnóza nebola súčasťou dát, ktorými sme disponovali, môžeme sa pri identifikácii hráčov s potenciálne problémovým hráčstvom spoliehať len na pomocné ukazovatele. K tejto úlohe sme pristúpili tak, že síce detegujeme anomálie v dátach, ale sme si vedomí toho, že nie všetky anomálie musia indikovať problém s gamblingom. Preto budeme korelovať výsledky AE modelu s týmito pomocnými ukazovateľmi:

**Tabuľka 4:** Zhoda medzi detegovanými množinami anomálií štyroch porovnávaných AE modelov.

Model	LSTM A	LSTM B	CNN	Transformer
LSTM A	100%			
LSTM B	65.1%	100%		
CNN AE	69.8%	61.1%	100%	
Transformer	41.3%	38.9%	38.9%	100%



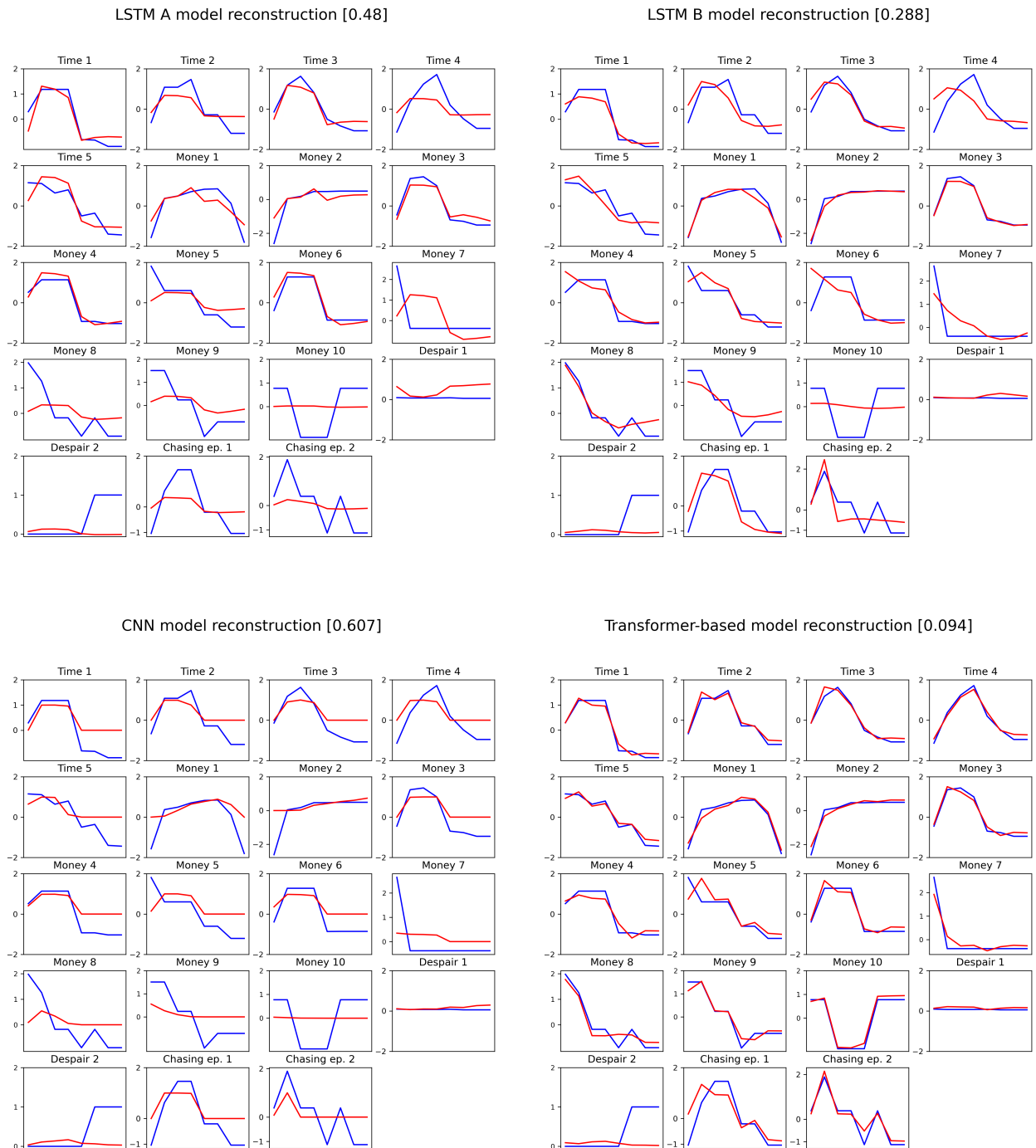
**Obr. 7:** Vennov diagram pre vizualizáciu prieniku detegovaných množín anomálií pre porovnávané AE modely. Hodnoty uvedené v diagrame predstavujú počet unikátnych hráčov detegovaných príslušným modelom/modelmi. Množina anomálií detegovaná každým modelom obsahuje 126 hráčov, čo odpovedá piatim percentám veľkosti testovacej množiny, keďže za anomálie považujeme pozorovania, ktorých skóre anomálie je väčšie alebo rovné než 95-ty percentil rekonštrukčnej chyby.

- Priemerný počet prihlásení v časovom okne.
- Priemerný počet výberov v časovom okne.
- Priemerný počet malých a častých výberov v časovom okne.
- Priemerný počet žiadostí o zmenu limitu výšky vkladov v časovom okne.
- Celkový počet epizód naháňania straty v období  $N$  časových okien.

Obrázok 9 zobrazuje koreláciu skóre anomálie s týmito pomocnými ukazovateľmi. Každý z grafov obsahuje desať stĺpcov, pričom každý stĺpec reprezentuje jeden decil pozorovaní (t.j. každý stĺpec zobrazuje 10% pozorovaní zoradených podľa hodnoty skóre anomálie). Farby stĺpcov predstavujú hodnotu/kategóriu príslušného pomocného ukazovateľa.

Na každom grafe je možné pozorovať zreteľné vzorce v správaní hráčov – hráči s vysokou hodnotou skóre anomálie majú tendenciu mať tiež vysoké hodnoty všetkých skúmaných pomocných ukazovateľov. Vysoká frekvencia prihlásení do online kasína je úmerná vysokému skóre anomálie, pričom viac než polovica hráčov v poslednom decile rekonštrukčnej chyby má priemerný počet prihlásení v časovom okne väčší než 50. Niečo podobné pozorujeme aj pre priemerný počet výberov v časovom okne. Hráči s nízkou hodnotou skóre nemajú žiadny, alebo majú veľmi malý priemerný počet výberov a zároveň viac než štvrtina hráčov v poslednom decile skóre má v priemere až dva a viac výberov v časovom okne. Ďalším skúmaným pomocným ukazovateľom je počet malých a častých výberov. Väčšina hráčov, ktorí majú v histórii aspoň jednu takúto udalosť, sú v 10% hráčov s najvyšším MSE. Pri analýze ďalšieho ukazovateľa, konkrétne počtu žiadostí o zmenu limitu vkladu, pozorujeme už menej výrazný trend. Je evidentné, že hráči v prvých piatich deciloch nemajú vo všeobecnosti žiadne zmeny limitu výberov (až na pár výnimiek), čím však skóre anomálie rastie, tým sa frekvencia zmeny limitu výberov u hráčov zvyšuje. Posledným pomocným ukazovateľom zobrazeným na grafoch je počet epizód naháňania straty. Opäť môžeme pozorovať zvyšujúcu sa frekvenciu týchto udalostí úmernú skóre anomálie. Viac než polovica hráčov v poslednom decile má aspoň jednu takúto epizódu v časovom okne.

Ak vykonáme superpozíciu týchto grafov za účelom zistenia, aký podiel hráčov spĺňa viacero pomocných ukazovateľov, dospejeme k nasledovnému pozorovaniu: 98.6% hráčov v posledných piatich percentiloch skóre anomálie spĺňa aspoň jeden pomocný ukazovateľ a 77.3% spĺňa aspoň tri ukazovatele. Analýzou správania hráčov v posledných dvoch percentiloch skóre ano-



**Obr. 8:** Porovnanie predikčnej schopnosti AE modelov. Všetky modely rekonštruovali rovnaké pozorovanie, pochádzajúce z testovacej množiny. Modrá čiara predstavuje vstupné dáta, červená rekonštrukciu získanú pomocou príslušného AE modelu. Číslo uvedené v nadpise grafov predstavuje skóre anomálie pre danú vzorku dát.

málie zistíme, že takmer 90% z nich spĺňa aspoň tri ukazovatele. Hranice, ktoré boli použité na výpočet týchto podielov sú:  $\geq 1$  epizóda naháňania straty,  $\geq 1$  žiadosť o zmenu limitu,  $\geq 1$  malý a častý výber,  $\geq 31$  prihlásení a  $\geq 1.25$  výberu v priemere počas jedného časového

vého okna.

Aj keď vek a pohlavie neboli použité ako prediktory v skúmaných modeloch, analyzovali sme aj demografický profil hráčov v spojitosti s ich skóre anomálie. Z výsledkov prezentovaných na Obrázku 10 môžeme konštatovať, že neexistuje



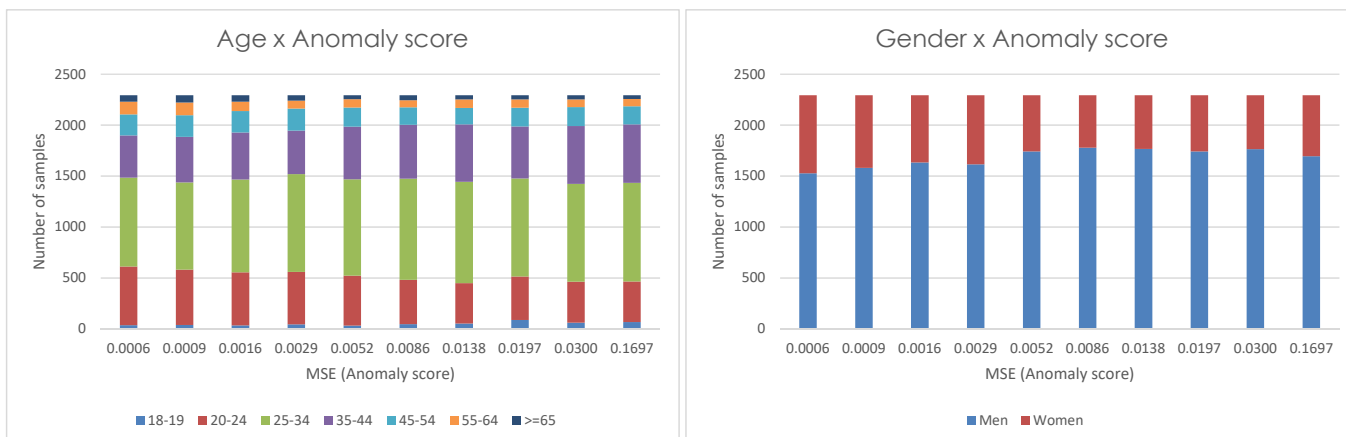
**Obr. 9:** Každý stĺpec v grafoch predstavuje jeden decil skóre anomálie (MSE). Farby predstavujú kategórie príslušných pomocných ukazovateľov s hodnotami kategórií špecifikovanými v legende.

jasný vzťah medzi vekom a skóre anomálie, pretože všetky vekové kategórie sú v každom decile skóre zastúpené približne rovnako. Z hľadiska korelácie pohlavia, pozorujeme nepatrný vzťah medzi touto veličinou a rekonštrukčnou chybou – ženy majú o niečo výraznejšie zastúpenie v prvých deciloch skóre s nižšou rekonštrukčnou chybou.

## Záver

V práci sme úspešne aplikovali model autoenkódera (AE) založený na transformeroch na problém detekcie anomálií v dátovej sade hráčov online kasína. Cieľom bolo detegovať hráčov

s patologickým hráčstvom metódami "bez učiteľa". Zo vstupných dát bolo odvodených 19 prvkov deskriptora vo forme časových radov, ktoré odrážajú správanie hráčov v kontexte "času", "peňazí" a "zúfalstva". Porovnali sme detekčnú schopnosť tejto architektúry s ďalšími tromi AE architektúrami založenými na LSTM a konvulčnými vrstvami, pričom sme zistili, že architektúra modelu založená na transformeroch dosiahla najlepšie výsledky zo všetkých skúmaných modelov v zmysle najlepšej rekonštrukčnej schopnosti. Výsledky tohto modelu tiež vykazujú vysokú koreláciu s pomocnými ukazovateľmi ako počet prihlásení hráča, počet výberov, počet epizód naháňania straty a ďalšími, ktoré sú často spo-



**Obr. 10:** Ľavý graf zobrazuje vzťah medzi vekom hráčov a ich skóre anomálie. Pravý graf obsahuje vizualizáciu vzťahu medzi pohlavím hráčov a ich skóre anomálie. Každý stĺpec predstavuje jeden decil rekonštrukčnej chyby (MSE), t.j. 10% hráčov zoradených podľa rekonštrukčnej chyby. Farby predstavujú kategórie veku/pohlavia špecifikované v legende.

mínané v literatúre vo vzťahu k patologickému hraniu. Tento súlad "skóre anomálie", ktoré odráža pravdepodobnosť, že dané pozorovanie sa vymyká priemeru, s pomocnými ukazovateľmi nám dovoľuje porozumieť modelu a jeho efektívnosti v detekcii hráčov s potenciálnym patologickým hráčstvom. Aj keď tieto pomocné ukazovatele boli použité ako prediktory v modeli, navrhujeme ich používať spolu s rekonštrukčnou chybou pri detekcii potenciálne problémových hráčov, aby sa redukovala falošná pozitivita predikcií, nakoľko nie všetky anomálie musia byť spojené s patologickým hráčstvom.

## Podakovanie

Tento projekt vznikol s podporou Spoločnej iniciatívy európskeho vysokovýkonného počítania (EuroHPC JU), číslo grantu 101101903. EuroHPC JU je podporované Európskou úniou prostredníctvom programu Digitálna Európa a členských krajín Nemecko, Bulharsko, Rakúsko, Chorvátsko, Cyprus, Česká republika, Dánsko, Estónsko, Fínsko, Grécko, Maďarsko, Írsko, Taliansko, Lotyšsko, Litva, Poľsko, Portugalsko, Rumunsko, Slovinsko, Španielsko, Švédsko, Francúzsko, Holandsko, Belgicko, Luxembursko, Slovensko, Nórsko, Švajčiarsko, Turecko, Republika Severné Macedónsko, Island, Čierna Hora a Srbsko.

## Referencie

- [1] Alex Blaszczynski a Lia Nower. "A Pathways Model of Problem and Pathological Gambling". In: *Addiction (Abingdon, England)* 97 (jún 2002), s. 487–99. DOI: 10.1046/j.1360-0443.2002.00015.x.
- [2] National Research Council. *Pathological Gambling: A Critical Review*. Washington, DC: The National Academies Press, 1999. ISBN: 978-0-309-06571-9. DOI: 10.17226/6329. URL: <https://nap.nationalacademies.org/catalog/6329/pathological-gambling-a-critical-review>.
- [3] Luke Clark et al. "Pathological Choice: The Neuroscience of Gambling and Gambling Addiction". In: *Journal of Neuroscience* 33.45 (2013), s. 17617–17623. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.3231-13.2013. eprint: <https://www.jneurosci.org/content/33/45/17617.full.pdf>. URL: <https://www.jneurosci.org/content/33/45/17617>.
- [4] Xiaolei Deng, Tilman Lesch a Luke Clark. "Applying Data Science to Behavioral Analysis of Online Gambling". In: *Current Addiction Reports* 6 (sept. 2019). DOI: 10.1007/s40429-019-00269-9.
- [5] Gaëlle Challet-Bouju et al. "Modeling Early Gambling Behavior Using Indica-

- tors from Online Lottery Gambling Tracking Data: Longitudinal Analysis". In: *J Med Internet Res* 22.8 (aug. 2020), e17675. ISSN: 1438-8871. DOI: 10.2196/17675. URL: <http://www.ncbi.nlm.nih.gov/pubmed/32254041>.
- [6] Michael Auer et al. "Development of the Online Problem Gaming Behavior Index: A New Scale Based on Actual Problem Gambling Behavior Rather Than the Consequences of it". In: *Evaluation & the Health Professions* (máj 2023), s. 1–12. DOI: 10.1177/01632787231179460.
- [7] Deepanshi Seth et al. "A Deep Learning Framework for Ensuring Responsible Play in Skill-based Cash Gaming". In: *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)* (2020), s. 454–459.
- [8] Leonardo Motta Perazzo Lannes. "Unsupervised Learning Applied to the Segmentation of Users of Online Gambling Platforms in Portugal". In: *NIMS - Dissertações de Mestrado em Ciência de Dados e Métodos Analíticos Avançados (Data Science and Advanced Analytics)* (nov. 2021).
- [9] Zhaomin Chen et al. "Autoencoder-based network anomaly detection". In: *2018 Wireless Telecommunications Symposium (WTS)*. 2018, s. 1–5. DOI: 10.1109/WTS.2018.8363930.
- [10] Nikolay Laptev, Saeed Amizadeh a Ian Flint. "Generic and Scalable Framework for Automated Time-Series Anomaly Detection". In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '15. Sydney, NSW, Australia: Association for Computing Machinery, 2015, s. 1939–1947. ISBN: 9781450336642. DOI: 10.1145/2783258.2788611. URL: <https://doi.org/10.1145/2783258.2788611>.
- [11] Saroj Gopali et al. *A Comparative Study of Detecting Anomalies in Time Series Data Using LSTM and TCN Models*. 2021. arXiv: 2112.09293 [cs.LG].
- [12] Ashish Vaswani et al. *Attention Is All You Need*. 2017. arXiv: 1706.03762 [cs.CL].
- [13] Pedro Lara-Benítez et al. "Evaluation of the Transformer Architecture for Univariate Time Series Forecasting". In: *Advances in Artificial Intelligence*. Ed. Enrique Alba et al. Cham: Springer International Publishing, 2021, s. 106–115. ISBN: 978-3-030-85713-4.
- [14] Alireza Makhzani et al. "Adversarial Autoencoders". In: *CoRR* abs/1511.05644 (2015). arXiv: 1511.05644. URL: <http://arxiv.org/abs/1511.05644>.
- [15] Thomas Schlegl et al. "Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery". In: *CoRR* abs/1703.05921 (2017). arXiv: 1703.05921. URL: <http://arxiv.org/abs/1703.05921>.
- [16] Xiaoran Chen a Ender Konukoglu. "Unsupervised Detection of Lesions in Brain MRI using constrained adversarial autoencoders". In: *CoRR* abs/1806.04972 (2018). arXiv: 1806.04972. URL: <http://arxiv.org/abs/1806.04972>.
- [17] Donny Hanz. *Data Exploration with Adversarial Autoencoders*. Towards Data Science. Júl 2019. URL: <https://towardsdatascience.com/data-exploration-with-adversarial-autoencoders-311a4e1f271b>.
- [18] Theodoros Ntakouris. *Time series Classification with Transformer*. 2021. URL: [https://keras.io/examples/timeseries/timeseries\\_classification\\_transformer/](https://keras.io/examples/timeseries/timeseries_classification_transformer/).