

Eötvös Loránd Tudományegyetem

Budapesti Corvinus Egyetem

Kópé Boglárka Flóra

HALANDÓSÁG ELŐREJELZÉSE GÉPI TANULÁSI MÓDSZEREKKEL

Szakdolgozat

Biztosítási és pénzügyi matematika mesterszak

Aktuárius szakirány

Témavezető(k): Szentkereszti Gábor, Dr. Vékás Péter



Budapest, 2024

Tartalom

1	Bevezetés.....	3
2	Szakirodalmi áttekintés	4
3	A Lee-Carter modell bemutatása.....	8
4	Gépi tanulási modellek bemutatása.....	10
4.1	Gépi tanulás.....	10
4.2	Mesterséges neurális hálózatok.....	10
4.3	Előrecsatolt neurális hálózatok.....	12
4.4	Optimalizáció	12
4.5	Visszacsatolt neurális hálózatok.....	15
4.6	Hosszú-rövidtávú emlékezet	16
4.7	Kapuzott visszacsatolt cella	18
5	Adatok bemutatása	20
6	A Lee-Carter modell eredményeinek ismertetése	21
7	A neurális hálózatok eredményeinek ismertetése.....	23
7.1	Neurális hálózatok hiperparaméter-optimalizálása	23
7.2	A hiba nagyságának alakulása.....	24
7.3	Becsült mortalitási ráták alakulása.....	26
8	A modellek eredményeinek összevetése	27
9	Előrejelzések jósága	31
9.1	Az élettartam-kockázatról	31
9.2	Várható hátralévő élettartam és életjáradék számítása.....	31
9.3	Az életjáradékokkal kapcsolatos eredmények.....	33
10	Összefoglalás.....	39

Ábrajegyzék

1. ábra: Neuron reprezentációja (Roell, 2017)	11
2. ábra: Hátraterjesztés szemléltetése (Szemenyei, 2021).....	14
3. ábra: Visszacsatolt architektúra felépítése (bal) és időbeli kibontása (jobb) (Szemenyei, 2021)	16
4. ábra: Az LSTM cella struktúrája (Szemenyei, 2021).....	17
5. ábra: GRU cella struktúrája, (Olah, 2015)	18
6. ábra: Valós logaritmált halandósági ráta életkoronként és naptári évenként.....	21
7. ábra: Lee-Carter modell paraméterei.....	22
8. ábra: MSE értékének alakulása modelltípusonként	25
9. ábra: Logaritmált halandósági ráta életkoronként (2000)	26
10. ábra: Halandósági ráta életkoronként (2000)	26
11. ábra: Logaritmált halandósági ráta (2020)	27
12. ábra: Halandósági ráta (2020)	28
13. ábra: A Lee-Carter és az LSTM struktúrájú neurális hálós modell reziduálisai.....	28
14. ábra: Halandósági ráta életkoronként	29
15. ábra: Halandósági ráta alakulása a 65 éves életkor feletti korosztályokban	35
16. ábra: 65 évesek halálozási valószínűségeinek előrejelzése	37
1. táblázat: Statisztikai és gépi tanulási módszerek különbségei (Vékás, Bevezetés a mesterséges neurális hálózatokba, 2022)	10
2. táblázat: Optimális hiperparaméterek.....	24
3. táblázat: 10 legkisebb MSE-jű modell átlaga típusonként	25
4. táblázat: MSE modellenként	27
5. táblázat: Összehasonlítás: a 65 éves korban várható hátralevő élettartam és az életjáradék egyszeri nettó díja (Májér & Kovács, 2011), (Vékás, Nyugdíjcélú életjáradékok élettartamkockázata az általánosított korcsoport-időszak-kohorsz modellkeretben, 2017).....	34

1 Bevezetés

A gépi tanulás és mesterséges intelligencia térnyerése az adatelemzés területén megkérdőjelezhetetlen. Ennek lehetőségét nemcsak az informatikai fejlődés, hanem a hatalmas mennyiségben rendelkezésre álló információ (big data) és adatbázisok elérhetősége teremtette meg. Annak ellenére, hogy gép tanulás és mesterséges intelligencia alapjait nem a közelmúltban fektették le szükséges volt, hogy megfelelő idő elteljen ahhoz, hogy az informatikai technológiák és a rendelkezésre álló információ mennyisége olyan szintre eljusson, ahol már alkalmazhatók ezen módszertanok.

Az idő múlása és a központi adminisztráció fejlődése lehetőséget adott nemcsak a keresztmetszeti, de hosszú idősoros gazdasági adatok elérhetőségére is. Ezen adatok elemzésére pedig a klasszikus statisztikai módszerek mellett alkalmazhatók lettek az adattudomány módszertanai. Ma még vitatható, hogy mely tudományág módszerei a megfelelőbbek a gazdasági adatok elemzésére, azonban az adatmennyiség gyarapodásával a pontosságbeli különbségek elhalványodni látszanak.

Ezen megfontolásból merült fel a gépi tanulás aktuáriusi területeken való alkalmazhatósága. Szakdolgozatomban az aktuáriusi gyakorlatban gyakran alkalmazott Lee-Carter halandóság-előrejelző modell jóságát kívánom összehasonlítani a gépi tanulás egyik népszerű módszertanának, a mesterséges neurális hálózatok előrejelzéseinek jóságával. Céлом megmutatni, hogy a mesterséges neurális hálózatok képesek éppen olyan jó – esetleg jobb – előrejelzéseket előállítani, mint a Lee-Carter modell. További vizsgálat alá vetem a neurális hálózatok hosszútávú előrejelzési képességeit és azok élettartam-kockázat számszerűsítésére való alkalmazhatóságát.

A dolgozatban bemutatásra kerülnek a releváns szakirodalmak, illetve az alkalmazott módszertanok és adatok. Ezt követően az egyes modellek eredményeinek ismertetésére és ezek összehasonlítására kerül sor. Végül a legjobb gépi tanulási modell hosszútávú előrejelzéseinek alkalmazhatóságát tesztelem az élettartam-kockázat számszerűsítésére.

2 Szakirodalmi áttekintés

Ebben a fejezetben a releváns szakirodalmak kerülnek bemutatásra.

Lee és Carter egy olyan módszertant dolgozott ki, amelyet ma napig előszeretettel alkalmaznak a halandóság előrejelzésére, annak erős statisztikai megalapozottsága miatt. Pontosan emiatt vált az egyik legelfogadottabb módszertanná a területen. A tanulmány, amelyben a már csak Lee-Carter modell néven emlegetett módszertant mutatták be alapvetően a halandóság és a születéskor várható élettartam javulásának kérdését járja körül. A vizsgálatot az Egyesült Államok lakosságára végezték el az 1933-1987-es időszakot tekintve. A paraméterek és változók definiálása után felírták a modellt, melyet aztán illesztettek a vizsgált korszpecifikus halandósági táblákra. Hangsúlyozzák, hogy nem egy klasszikus regressziós modell lévén nem lehet azt az általános regressziós módszerekkel illeszteni, mert nincsenek - a regressziónál megszokott - magyarázó változók. Emiatt szinguláris érték felbontással határozták meg a legkisebb négyzetes megoldást. A modellt életkoronként illesztik, így könnyen megállapítható az életkor logaritmált halálozási rátákra gyakorolt hatása. A modellen belül is a néphalandóságot leíró paramétert külön, standard statisztikai módszerekkel becsülték, eltolásos véletlen bolyongást alkalmazva. Ezen idősort már hosszabb időszakon vizsgálják, 1900-1989 között. Az 1918-as influenza járványra külön kontrollálnak, hogy annak torzító hatása ne jelenjen meg az eredményekben. A logaritmálás miatt a különböző életkorokhoz tartozó halandóság ugyan eltérő mértékben, de minden esetben exponenciális sebességgel fog csökkenni. Az előrejelzések alapján születéskor várható élettartam alakulását is megállapították, 2065-re 86,05-re becsülték a születéskor várható élettartamot. (Lee & Carter, 1992)

Hazai viszonylatban is készült tanulmány a halandóság modellezésére. Májer és Kovács cikkükben az élettartam-kockázat nyugdíjrendszerre gyakorolt hatásáról írnak. Röviden bemutatják a magyar nyugdíjrendszert, illetve a nyugdíjrendszerek fajtáit. A tanulmány fő célja az időskori várható élettartam jövőbeli alakulásának modellezése. Definiálásra kerülnek a legfontosabb fogalmak, mint halálozási valószínűség, valamint az abból származtatott túlélési valószínűség és várható élettartam. Utóbbi megállapítására szükségük volt a halálozási valószínűségek előrejelzésére. Azzal a feltételezéssel éltek, hogy a halálozási valószínűségekben végbemenő változás egy véletlenszerű (sztochasztikus) folyamat, ugyanis ebből a sztochasztikus trendből adódik az ún. élettartam-kockázat. A nyugdíjkalkulációk során három kockázatot kell megfontolni: az egyéni halálozási, várható élettartam és modellkockázatot. Ezen kockázatok megszüntetésének lehetőségéről és módszereiről is

részletes leírást készítettek. A mortalitás modellezéséhez az előző cikkben bemutatott, Lee-Carter modell került alkalmazásra, annyi módosítással, hogy úgynevezett változsfaktort is használtak. A változsfaktor alkalmazásával akarták elkerülni, hogy egy hirtelen ugrás legyen az utolsó év megfigyelt és az első előre jelzett év mortalitási rátáinak értékei között. A modell eredményei alapján megállapították és előrejelezték a mortalitási ráták és jövőbeli várható élettartam nagyságát illetve az akörüli bizonytalanságot. A fókusz ebben az esetben a 65 év felettiekre került, hiszen a nyugdíjszámítás tekintetében ez a korcsoport a legfontosabb. A Lee-Carter modell eredményei alapján kétféleképpen számították ki a túlélési valószínűségeket és összehasonlították a kapott értékeket. Az első esetben a jövőbeli mortalitási szint trendjét ismertnek, a második esetben pedig egy valószínűségi változónak tekintették. A cikk alapvetően az egyre idősödő társadalom nyugdíj jellegű problémáit kívánta szemléltetni, annak aktualitását hangsúlyozva. (Májér & Kovács, 2011)

Vékás is elkészített egy elemzést a magyarországi halandóság és élettartam kockázat vizsgálatára. Kutatása során az általánosított korcsoport-időszak-kohorsz (GAPC, generalized age-period-cohort – általánosított kor-csoport-időszak-kohorsz) modellesaládot alkalmazta és eredményeit összevetette Májér István és Kovács Erzsébet cikkében született eredményekkel. A halandóság változásáról írva bemutatja az élettartam-kockázatot és az ennek figyelmen kívül hagyásából fakadó pontatlanságot is. Ezen pontatlanságok alatt a nyugdíjrendszerek, illetve járadéktermékeket értékesítő biztosítók bevételeinek és kiadásainak hibás kalkulációját értjük. Kiemeli, hogy 2016. január 1-jétől, a Szolvencia II EU keretirányelv is előírja kezelésének előtérbe helyezését. Bemutatja a már említett GAPC-modellkeretet: a korcsoport- és időszakspecifikus változókat és az ezekre vonatkozó feltételezéseket, illetve a modell ún. szisztematikus komponensét, amely a korcsoport, időszak és kohorsz függvényeként írható fel. A modell paramétereinek értelmezését is átfogóan bemutatja. A GAPC-modellcsalád nevezetes modelljeiről is ír, mint a Poisson-féle Lee-Carter- (LC-), A Renshaw-Haberman- (RH-), a korcsoport-időszak-kohorsz (APC-), a Cairns-Blake-Dowd- (CBD-) és a Plat-modell. Ezen modellekre el is készítette maximum likelihood módszer segítségével a becsléseket, majd modellek performancájának összevetése után a Cairns-Blake-Dowd- (CBD-) modellt ítélte meg a legjobbnak. A választott modell segítségével elvégzi a nyugdíjkorhatáron várható hátralevő élettartam és a nyugdíjcélú életjáradékok egyszeri nettó díjának becslését. Számításai alapján keresztmetszeti szemléletben vett 65 éves korban várható élettartam 2006-2014 között 1,08 évvel és a nyugdíjcélú életjáradék díja pedig 1,85 Ft-tal nőtt a Májér István és Dr. Kovács

Erzsébet által számolthoz képest. (Vékás, Nyugdíjcélú életjáradékok élettartam-kockázata az általánosított korcsoport-időszak-kohorsz modellkeretben, 2017)

A gépi tanulás és mesterséges intelligencia alkalmazásainak kutatása nagy népszerűségnek örvend az utóbbi időszakban, szakdolgozatomhoz kapcsolódóan két a témába írt cikket szeretnék részletesebben bemutatni.

Nigri, Levantesi, Marino, Scognamiglio és Perla tanulmányukban a halandóságot Lee-Carter modellbe integrált gépi tanulás alkalmazásával becsülték meg és eredményeiket összehasonlították az eredeti Lee-Carter modell eredményeivel. A Lee-Carter modell két paraméterének ARIMA becslésének korlátaira felhívva a figyelmet, egy olyan alternatív megközelítést javasoltak, amely a nemlinearitást is jól tudja kezelni. Ez hosszú rövidtávú emlékezet (LSTM, Long Short-Term Memory) módszere, amely a visszacsatolós neurális hálók (RNN, Recurrent Neural Networks) egy speciális fajtája. Annak ellenére, hogy a gépi tanulás alkalmazása iránti érdeklődés nagyban megnőtt az utóbbi években bizonyos tudományterületeken, illetve a demográfiai kutatásokban még mindig nem vált népszerűvé, köszönhetően annak, hogy nehezen értelmezhető. Az írók kiemelik, hogy a népszerűtlenséget okozhatja az is, hogy a gépi tanulás egy adat-orientált módszer, azonban a demográfusok jobban preferálják hipotézis-alapú megközelítéseket. Az írók úgy gondolják, hogy a halandóság dinamikáját a neurális hálózatok az eddig használt megközelítéseknél jobban meg tudják ragadni. Részletesen bemutatják a neurális hálózatok, működését és limitációit az idősoros előrejelzésekre. Majd ezen limitációk feloldására bevezetik a visszacsatolós neurális hálókat, illetve a hosszú rövidtávú emlékezetet. Az elemzést hat országra végezték el: Ausztrália, Dánia, Olaszország, Spanyolország, Egyesült Államok és Japán. Minden esetben az LSTM integrált modellek teljesítettek jobban. (Nigri, Levantesi, Marino, Scognamiglio, & Perla, 2019)

Richman és Wüthrich más megközelítéssel éltek a gépi tanulás a halandóság előrejelzésekre való alkalmazásánál. Az (Nigri, Levantesi, Marino, Scognamiglio, & Perla, 2019) tanulmányában említett módszertant nem integrálva, közvetett módon, hanem közvetlenül használták a halandóság előrejelzésére. Tehát a visszacsatolós neurális hálók különböző változatait alkalmazták, melyeket a már a sokat emlegetett, Lee-Carter modell teljesítményével hasonlítottak össze. A svájci halandósági rátákat tanulmányozták és ezek előrejelzésére készítettek mesterséges neurális hálós modelleket. A visszacsatolós neurális hálók két speciális változatát mutatták be részletesen, mint alkalmazott módszertant: a már korábban is említett LSTM architektúrát és a kapuzott visszacsatolt cella (GRU, Gated Recurrent Unit). A hiperparaméterek megfelelő megválasztásának tárgyalása után bemutatták továbbá az időben

elosztott LSTM hálózatot (Time-distributed LSTM), amely vektor formában prediktálja a kimeneteket. Elemzésük alapján az LSTM és a GRU hálózatokkal pontosabb eredményeket lehet elérni, mint a Lee-Carter-modell alkalmazása esetén. A hálózatok stabilitását kérdésesnek tartják, mert a különböző tanítások során eltérő eredményeket generáltak, így ensemble technikákkal való döntést javasolják. (Szentkereszti & Vékás, 2022) (Richman & Wüthrich, 2019)

A gépi tanulás, mint halandóság előrejelző technika alkalmazására Szentkereszti és Vékás cikkében is sor került. Magyar halandósági rátákat jeleztek előre visszacsatolt neurális hálók alkalmazásával. Tanulmányukban továbbá röviden kifejtették a gépi tanulási módszerek lehetőségeit az aktuáriusi területen. A cikk központi módszertana az LSTM architektúra. Az összehasonlítást ebben az esetben is a Lee-Carter modell eredményeivel végezték. Az előző cikkekhez hasonlóan, arra az eredményre jutottak, hogy pontosabb értékeket lehet elérni a visszacsatolt neurális hálós technikák alkalmazásával. (Szentkereszti & Vékás, 2022)

3 A Lee-Carter modell bemutatása

A dolgozatban az alkalmazott gépi tanulási modellek minden esetben a Lee-Carter halandóság-előrejelző modellel kerülnek összehasonlításra, így ebben a fejezetben ennek módszertanát kívánom bemutatni. Először a halandóság-előrejelzés legalapvetőbb fogalmai, majd maga a Lee-Carter modell kerül bemutatására.

A halandóság modellezésében leggyakrabban használt statisztikai mutató a *halandósági ráta* vagy *halandósági arányszám*, amely egy választott időszak és populáció vonatkozásában értelmezhető, és az adott időszak során az adott populációban bekövetkezett halálozások számának a populáció létszámához viszonyított arányaként számítható ki. Formulával:

$$m = \frac{D}{E}$$

ahol m a halandósági ráta, $D \in \mathbb{N}$ a vizsgált időszakban elhunytak száma, $E > 0$ pedig a vizsgált populáció létszáma. Jelen dolgozatban a vizsgált időszak alatt élő egyének átlagos létszámát, úgynevezett *központi kitétséget* fogok érteni, jele: E^c . Ez a vizsgált időszak kezdetén életben lévő egyénekre a vizsgált időszakban megélt egyéni időmennyiségeket összegezve számítható ki. Központi kitétség alkalmazása esetén *központi halandósági rátáról* beszélhetünk. További kiegészítéseket is tettem a halandósági rátára. Életkortól $x \in \{1, 2, \dots, X\}$ és naptári évtől $t \in \{1, 2, \dots, T\}$ is függővé tettem. Így a halandósági ráta végleges formulája:

$$m_{xt} = \frac{D_{xt}}{E_{xt}^c}$$

A Lee-Carter modell a következő egyenlettel írja le a halandóság alakulását:

$$\ln(m_{xt}) = a_x + b_x k_t + \varepsilon_{xt}$$

A paraméterek értelmezése a következő:

- $\ln(m_{xt})$: logaritmált halandósági ráta.
- a_x : átlagos logaritmált halandósági ráta értéke az adott korcsoportban.
- k_t : mortalitási index, az egyetlen időtől függő paraméter, amely a mortalitás időbeli változását testesíti meg.
- b_x : életkorfüggő érzékenység, azaz a logaritmált halandósági ráta érzékenysége a mortalitási index változásának hatására
- ε_{xt} : modell véletlen hibtagja, gyakori feltevés, hogy független normális eloszlású, azonosan 0 várható értékkel és $\sigma^2 > 0$ varianciával.

Identifikációs problémát okozhat, hogy a paraméterek nem egyértelműek, így szükséges a következő megkötéseket alkalmazni a paraméterbecslés során:

$$\sum_{x=1}^X b_x = 1$$

$$\sum_{t=1}^T k_t = 0.$$

A paraméterek maximum likelihood becslése a fenti egyenlet és megkötések melletti log-likelihood függvény maximalizálásával érhető el:

$$\Theta = \mathbb{R}^{2X+T} \times \mathbb{R}^+$$

$$\boldsymbol{\theta}^T = (\mathbf{a}^T, \mathbf{b}^T, \mathbf{k}^T, \sigma) \in \Theta$$

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta} \in \Theta} \{\ell(\boldsymbol{\theta}) : \mathbf{1}^T \mathbf{b} = 1, \mathbf{1}^T \mathbf{k} = 0\}$$

$$l(\boldsymbol{\theta}) = \alpha - XT \ln(\sigma) - \frac{1}{2\sigma^2} \sum_{x=1}^X \sum_{t=1}^T (\ln(m_{xt}) - a_x - b_k k_t)^2$$

ahol $\alpha = -\frac{1}{2}XT \ln(2\pi)$ konstans. A felírásból látható, hogy $\boldsymbol{\xi}^T = (\mathbf{a}^T, \mathbf{b}^T, \mathbf{k}^T) \in \Xi = \mathbb{R}^{2X+T}$ paramétervektor maximum likelihood becslését tetszőleges $\sigma > 0$ esetén a feladat célfüggvényében szereplő eltérés-négyzetösszeg minimalizálása szolgáltatja:

$$\hat{\boldsymbol{\xi}} = \arg \min_{\boldsymbol{\xi} \in \Xi} \{SSE(\boldsymbol{\xi}) : \mathbf{1}^T \mathbf{b} = 1, \mathbf{1}^T \mathbf{k} = 0\}$$

$$SSE(\boldsymbol{\xi}) = \sum_{x=1}^X \sum_{t=1}^T (\ln(m_{xt}) - a_x - b_k k_t)^2$$

Bár a feladat jellegében hasonlít a klasszikus legkisebb négyzetek problémájához, megoldása nem végezhető a közönséges legkisebb négyzetek (OLS) módszerével, mivel a összefüggésben egyedül a bal oldalon szerepelnek ismert mennyiségek.

Az optimális megoldás az eltérés-négyzetösszeg redukált alakjának előállítás után szingulárisérték-felbontás (SVD) és az Eckart–Young–Mirsky approximációs tétel alkalmazásával megkapható. (Vékás, Az élettartam-kockázat modellezése, 2019) Utóbbi tételek és levezetések bemutatása nem célja a dolgozatnak.

4 Gépi tanulási modellek bemutatása

4.1 Gépi tanulás

A gépi tanulás az adattudomány (data science) egy fontos területe, illetve a mesterséges intelligencia egy ága, amely főként az adatok olyan jellegű feldolgozásával foglalkozik, ami az emberi tanulást imitálja. Ebben a tudományágban különböző statisztikai módszerekkel tanítják az algoritmusokat, hogy azok a megtanult mintázatok és összefüggések segítségével klasszifikálják, becsüljék vagy előrejelezzék az adatokat. Ennek ellenére mégsem a klasszikus statisztikai módszerek közé soroljuk. A legfőbb különbség a modellezési módszerek megközelítésben keresendő: a statisztikai módszerek hipotézisalapú, míg a gépi tanulási módszerek adat-orientáltak. További különbségek is megfigyelhetők, amit az alábbi táblázat foglal össze.

Szempont	Statisztika	Gépi tanulás
Eredet	Valószínűségszámítás	Számítástudomány
Fókusz	Oksági kapcsolatok	Predikció
Interpretáció	Nagyon fontos	“Fekete doboz”
Kor	Több száz év	Pár évtized
Előfeltételek	Vannak	Nincsenek
Tesztek (p-érték)	Vannak	Nincsenek

1. táblázat: Statisztikai és gépi tanulási módszerek különbségei (Vékás, Bevezetés a mesterséges neurális hálózatokba, 2022)

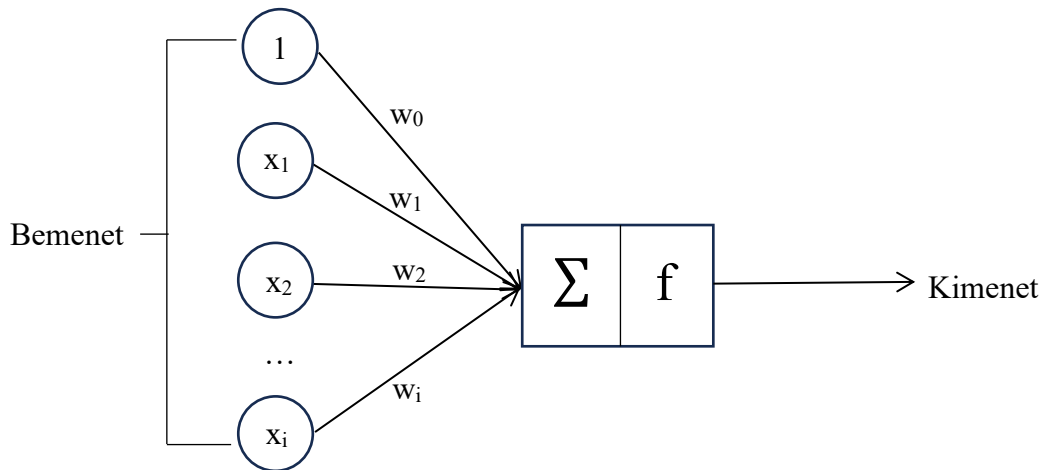
A táblázat alapján további magyarázatra szorulhat, hogy a gépi tanulási módszerek interpretáció szempontjából miért tekinthetők „fekete doboznak”. Egyrészt a módszer lépésenkénti értelmezése nehézkes. Másrészt nem mutatja meg, hogy hogyan jutott a kapott eredményre, így azon túl, hogy az általunk biztosított bementek alapján – amiket belerakunk a „fekete dobozba” - ad egy kimenetet, a modell nem ad további magyarázatot. Éppen emiatt is esik a fókusz a predikcióra, nem pedig az okságra.

Továbbá, ami sok szempontból is nagy előnyt jelenthet, hogy nincsenek előfeltételek az adattal kapcsolatban, illetve ezek formális tesztelésére sincsen szükség.

4.2 Mesterséges neurális hálózatok

A mesterséges neurális hálózatok a gépi tanulás módszerek egyik legnépszerűbb fajtája. Az emberi agyban megtalálható neuronok felépítése inspirálta a módszertant. Azt jelenséget próbálja matematikai módszerekkel formalizálni, ahogyan egy biológiai neuron a kapott impulzusokat továbbítja a többi elágazásra. Ugyanis a neuron ezen kapott impulzusokat nem

egyenként, hanem összegezve, egyetlen új impulzusként küldi tovább az rendszer további elágazásaira. Ha a bemeneti impulzusokat magyarázó változóknak a kimeneti impulzust pedig az eredményváltozónak nevezzük, akkor ezt az alábbi ábrával lehetne megmutatni. (Roell, 2017)



1. ábra: Neuron reprezentációja (Roell, 2017)

Az x_i -k az egyes bemeneti változókat szemléltetik, amiket w_i súlyokkal összegzünk. Ezen összegre alkalmazzuk az f nemlineáris függvényt, amelyet aktivációs függvénynek nevezünk. Formalizálva:

$$\hat{y} = f(\mathbf{w}^T \mathbf{x}) = f\left(\sum_{i=0}^n w_i x_i\right), \text{ ahol } x_0 = 1$$

Ha f éppen az előjelfüggvény ($\text{sgn}(x)$), akkor épp a Rosenblatt-perceptront definiáltuk. (Szentkereszti & Vékás, 2022)

További példaként felírható a lineáris regresszió, mint neurális hálózat az alábbi módon:

$$\hat{y} = \sum_{i=0}^n w_i x_i + b$$

Ez egy egyetlen darab neuront tartalmazó háló, identitás aktivációs függvénnyel. (Vékás, Bevezetés a mesterséges neurális hálózatokba, 2022)

Ha egy hálózatban több neuron is szerepel, akkor azokat rétegekbe rendezzük: bemeneti, rejtett, illetve kimeneti rétegbe. Ha egynél több rejtett rétege is van a hálózatnak, akkor mély neurális hálózatról beszélünk. A neurális hálózatot gyakran irányított gráffal ábrázoljuk, ahogy ezt az 1. ábra is mutatja. (Altrichter, és mtsai., 2006) A csomópontok a neuronokat, a gráf élei a neuronok

közötti kapcsolatokat jelölik. Ha a hálózatot leíró gráf nem tartalmaz hurkot, akkor előreccsatolt neurális hálózatról beszélünk (Feed-Forward Neural Network), ellenkező esetben visszacsatolt neurális hálózatról (RNN - Recurrent Neural Network). (Szentkereszti & Vékás, 2022)

4.3 Előreccsatolt neurális hálózatok

Az előreccsatolt neurális hálózat alatt, az előzőekben megfogalmazott struktúrát fogom érteni. Fontos, hogy a legelterjedtebb, teljesen összekötött hálózatot alkalmaztam, amely azt jelenti, hogy minden réteg, minden neuronja kapcsolódik a következő réteg minden neuronjához. Univerzális approximációs tétel értelmében kellően sok neuronnal egy ilyen háló bármilyen folytonos függvényt tetszőleges pontossággal megközelít. (Vékás, Bevezetés a mesterséges neurális hálózatokba, 2022) Az előreccsatolt neurális hálózatot formalizálva az alábbi módon lehet felírni:

$$1. \text{ réteg: } f_1(\mathbf{W}_1 \mathbf{x} + \mathbf{b}_1)$$

$$2. \text{ réteg: } f_2(\mathbf{W}_2 f_1(\mathbf{W}_1 \mathbf{x} + \mathbf{b}_1) + \mathbf{b}_2)$$

...

ahol $\mathbf{W}_1, \mathbf{W}_2, \dots$ a súlymátrixokat, $\mathbf{b}_1, \mathbf{b}_2, \dots$ a konstansokat, illetve f_1, f_2, \dots az aktivációs függvényeket jelöli. Ezen többszörösen összetett függvények áttekinthetőbb felírására vezessük be z_i -t, $i = 1, 2, \dots, k$:

$$1. \text{ réteg: } z_1 = f_1(\mathbf{W}_1 \mathbf{x} + \mathbf{b}_1)$$

$$i. \text{ réteg: } z_i = f_i(\mathbf{W}_i z_{i-1} + \mathbf{b}_i)$$

...

$$\text{Kimeneti réteg: } \hat{y} = f_k(\mathbf{W}_k z_{k-1} + \mathbf{b}_k)$$

4.4 Optimalizáció

Az ilyen többretegű hálózatok tanítása felügyelt tanítás, ahol a hálózat kimenetén értelmezett hiba segítségével határozhatjuk meg a hibafüggvény paraméterfüggését. A hibafelület a tanítandó paraméterek folytonos, differenciálható függvénye, ami a gradiens alapú tanuló algoritmusok alkalmazását lehetővé teszi. Ezek közül a leggyakrabban alkalmazott hibavisszaterjesztés, avagy hátraterjesztésnek (backpropagation). (Altrichter, és mtsai., 2006)

A tanítás módszerének alapvető kérdése, hogy hogyan lehet a súlyok értékét úgy meghatározni, hogy az osztályozás minél pontosabb legyen, valamint, hogy milyen hibafüggvény segítségével

mérhető jól az algoritmus teljesítménye. Célszerű egyszerűen az elvárt és a becsült kimenet közti átlagos négyzetes hibát (MSE - Mean Squared Error) venni, amely regressziós problémák esetén a leggyakrabban használt hibafüggvény. (Szemenyei, 2021) Így a dolgozatban alkalmazott modellek hibafüggvényének is ezt, illetve - további mérési célokra - ennek gyökét (RMSE - Root Mean Squared Error) választottam.

$$MSE: L(y_i, \hat{y}_i) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$RMSE: L(y_i, \hat{y}_i) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

A teljes hibát (TL - Total Loss) akarjuk minimalizálni:

$$\min TL(\mathbf{W}) = \sum_{i=1}^n L(y_i, \hat{y}_i)$$

Ha kiszámoljuk a hibafüggvény súlyok szerinti deriváltját (gradiensét), akkor megkapjuk azt, hogy hogyan kellene a súlyoknak megváltozni ahhoz, hogy a hibafüggvény a lehető leggyorsabban növekedjen. Ha azonban a súlyokat a gradienssel ellenkező irányba változtatjuk, az a leggyorsabb csökkenés iránya lesz. Ezt a lépést egyfajta „fordított hegymászóként” ismételve egy idő után lokális minimumba jutunk. Mivel a teljes hibát szeretnénk minimalizálni, ezért minden egyes lépéskor ki kell értékelni az egész tanító adatbázis hibáját. Ez a gradiens módszer esetén csak sok lépés után konvergál. Éppen ezért célszerű felosztani a tanító adatot egyenlő méretű, véletlenszerűen kiválasztott részhalmazokra, minikötegekre (minibatch-ekre), és minden ilyen minikötegre elvégezni egy lépést. Ezt a módszert sztochasztikus gradiens módszernek (SGD – Stochastic Gradient Descent) nevezzük. A gradiens módszer formalizálva:

$$\mathbf{W}_{t+1} = \mathbf{W}_t - \lambda * \nabla TL(\mathbf{W}_t)$$

ahol $\lambda > 0$ a tanulási ráta, ami egy olyan hiperparaméter, ami nagymértékben befolyásolja a tanítás sebességét és minőségét. (Szemenyei, 2021)

A gradiens módszer egyik hátránya, hogy könnyen beragadhat lokális minimumokba. Ennek megoldására az inverz hegymászó ötletét le kell cserélnünk a hegyről leguruló szikla képére. Tehát a gradiens módszerhez egyfajta momentumot veszünk hozzá. Ezt a gyakorlatban úgy

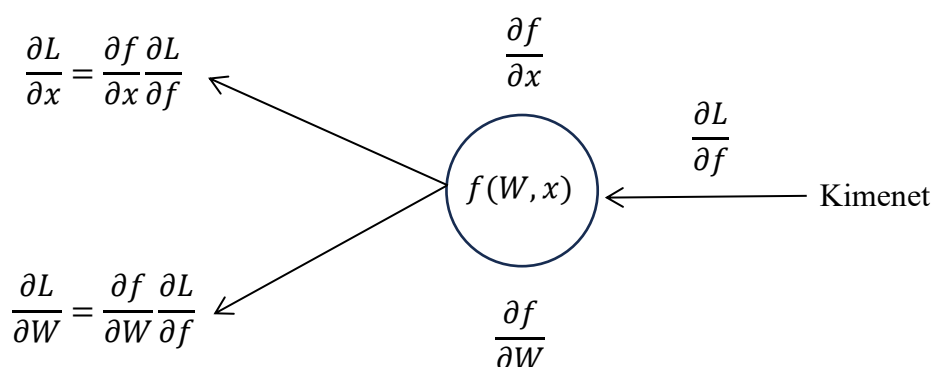
tesszük, hogy az adott időpontban elvégzett lépés a negatív gradiens iránya és az eggyel korábbi időpillanatban tett lépés súlyozott átlagából tevődik össze. (Szemenyei, 2021)

Az SGD egyik problémája, hogy a többdimenziós paraméterter minden irányát egyenértékűnek veszi, azonban előfordulhat, hogy a hibafüggvény az egyik irányban meredek, míg a másik irányban pont hogy lapos. Tehát az egyik irányba óvatosan kellene haladni, csökkenteni a lépésmértet, nehogy átugorjuk a minimumot, míg a másik irányban nagyobb lépésekkel lenne érdemes haladni. A kettőt egyszerre viszont nem lehet. Erre ad megoldást az adaptív tanulási ráta vagy gradiens skálázás. Alapelve az, hogy a gradiens nagyságát megjegyzi az egyes irányokban, a lépésméretet pedig ezek inverzével skálázza, így biztosítva az eltérő lépés méretet az egyes irányokban. (Szemenyei, 2021)

Az Adaptive Moment Estimation, Adam algoritmus (Kingma & Ba, 2015) beleépíti mind a minikötegek, a momentum és az adaptív tanulási ráta tulajdonságait a gradiens ereszkedés módszerébe. Ezért is vált e területen az egyik legnépszerűbb optimalizációs algoritmussá.

További kérdés lehet az optimalizáció kapcsán, hogy a súlyok szerinti deriváltakat milyen módon számoljuk. Egy neurális háló leírható, mint egy számítási gráf, ahol a gráf egyes csomópontjai deriválható függvényeket implementálnak. Amennyiben egy számítási gráfban ismerjük a bemeneteket, és az egyes csomópontok által implementált függvényeket, akkor az összes csomópont kimenetét ki tudjuk számolni. Ezt nevezzük az előreterjesztés műveletének. (Szemenyei, 2021)

Azonban, amennyiben ismerjük a csomópontok függvényeit, és a hibafüggvény deriváltját a csomópont kimenete szerint, akkor a deriválás láncszabályának segítségével könnyedén előállíthatjuk a bemenetek és súlyok szerinti deriváltakat. (Szemenyei, 2021) A 2. ábra is éppen ezt szemlélteti.



2. ábra: Hátraterjesztés szemléltetése (Szemenyei, 2021)

ahol $\frac{\partial f}{\partial x}$ és $\frac{\partial f}{\partial w}$ lokális gradiensek. Ezzel a módszerrel a hálón hátrafele haladva az összes bemenet és súly gradiensét meg tudjuk határozni. Ezt a műveletet hátraterjesztésnek (backpropagation) nevezzük. A gráf utolsó pontjának a kimenete maga a hibafüggvény. Egy mennyiség saját maga szerinti deriváltja pedig triviálisan 1. Ily módon tehát minden rendelkezésre áll a háló gradienseinek számolására. (Szemenyei, 2021)

4.5 Visszacsatolt neurális hálózatok

Az előre-csatolt neurális hálóknak memória eleme nincs, így nem igazán alkalmas idősoros adatok feldolgozására. Az ilyen idősoros adatok feldolgozásához viszont egy olyan új háló struktúrára lesz szükségünk, amely valamilyen belső állapottal is rendelkezik. Az ilyen hálózatokat visszacsatolt neurális hálózatoknak (RNN – Recurrent Neural Network) nevezzük. (Szemenyei, 2021)

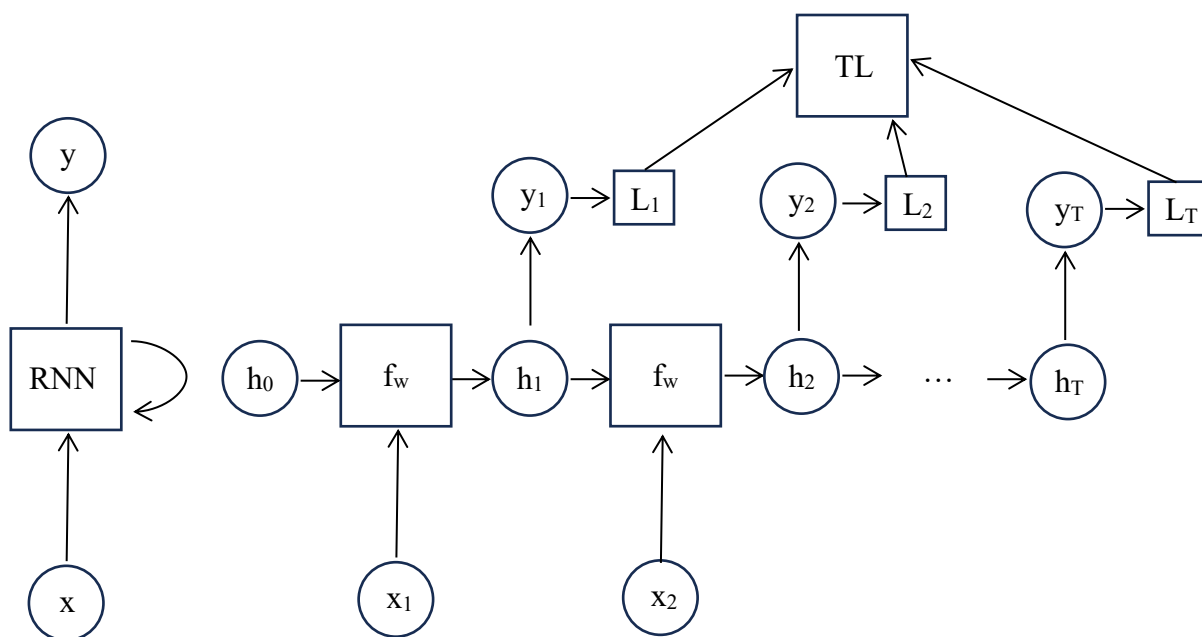
A visszacsatolt réteg működése során a belső állapotának aktuális értékét az aktuális bemenet és az egy lépéssel korábbi belső állapot értéke alapján számolja ki. A cella kimenete pedig a belső állapot imént kiszámolt aktuális értékétől függ. Formalizálva:

$$h_t = f(W_{hh}h_{t-1} + W_{xh}x_t + b_x)$$

$$y_t = W_{hy}h_t + b_y$$

ahol h a belső állapotot, t pedig az aktuális időpontot jelöli. (Szemenyei, 2021)

Egy RNN cella tulajdonképpen három lineáris réteg együtteseként adódik. A gradiens meghatározása azonban gondot jelenthet kezdetben. Probléma ugyanis, hogy a backpropagation módszere visszacsatolt neurális hálózatok esetén nem működik. Ez azonban könnyen kiküszöbölhető: egy visszacsatolt háló ugyanis átalakítható egy hagyományos előre-csatolt hálóvá az időben történő kibontás műveletével. Ez azt jelenti, hogy az egyetlen RNN réteg különböző időpontokban felvett állapotára úgy tekintünk, mint egy hagyományos háló egymást követő rétegeire. (Szemenyei, 2021)

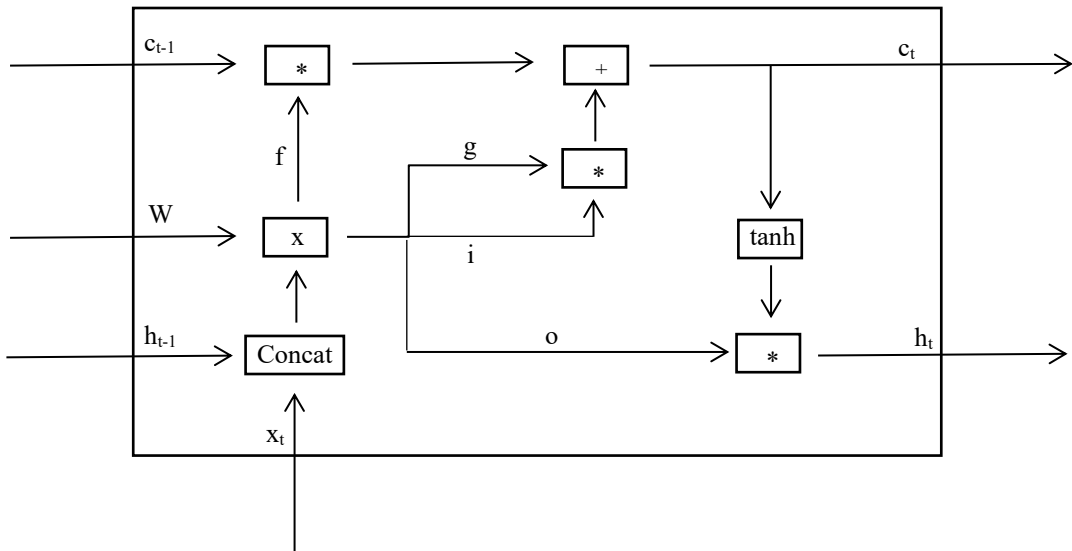


3. ábra: Visszacsatolt architektúra felépítése (bal) és időbeli kibontása (jobb) (Szemenyei, 2021)

Mivel egy RNN cellának minden időpontban van kimenete és hibája, ezért a kibontott háló minden rétegéhez fog tartozni egy kimenet és egy hiba, melyeknek összege éppen a teljes hiba. Innentől a hibavisszaterjesztés eljárás már minden további nélkül használható. Két fontos különbség adódik azonban az előrecsatolt hálókhoz képest. Egyrészt, ahogy haladunk előre az időben a kibontott háló mérete egyre nő – a végtelenségig -, ami a tanítás lassulásával jár. A megfelelő működéshez és tanításhoz nem szükséges a végtelenségig emlékezni a múltbeli bemenetekre. Éppen ezért a kibontás során a háló maximális méretét korlátozzuk és a legrégebbi réteget és bemenetet töröljük a kibontott hálóból.

4.6 Hosszú-rövidtávú emlékezet

Másrészt az időben a kibontott háló esetén minden réteg súlymátrixa azonos (hiszen valójában egyetlen visszacsatolt rétegről van szó). Ez azt jelenti, hogy amikor a láncszabály segítségével a deriváltakat előállítjuk, akkor az egyes rétegek deriváltjainak egy hosszú szorzatát kell kiszámolnunk, ahol jelen esetben egy hatványról beszélhetünk, hiszen a szorzat minden eleme azonos. Bármely szám (vagy mátrix) sokadik hatványa vagy nulla vagy végtelen, kivéve, ha az a szám pontosan 1. Így tehát egy visszacsatolt cella gradiensei könnyedén eltűnnek, vagy „felrobbannak”, ami a tanítást ellehetetleníti. Ezen probléma kiküszöbölésére egy olyan architektúra megalkotására volt szükség, ahol a belső állapot aktuális és egy lépéssel korábbi állapota közti derivált nagyjából egy. Éppen ilyen struktúrájú hosszú-rövidtávú emlékezet (LSTM - Long Short-Term Memory) cella. (Hochreiter & Schmidhuber, 1997.) Az LSTM az RNN cellához képest megfelelően hosszú ideig képes emlékezni. (Szemenyei, 2021)



4. ábra: Az LSTM cella struktúrája (Szemenyei, 2021)

ahol * az elemenkénti (Hadamard-) szorzat. Az LSTM cella négy aktivációs függvényt is tartalmazó egységből áll, melyeket kapuknak nevezünk. Ezen négy aktivációs függvény esetén három esetben a szigmoid és egy esetben pedig hiperbolikus tangens függvényt használunk:

$$\text{Sigmoid: } \sigma(x) = \frac{1}{1+e^{-x}}$$

$$\text{Hiperbolikus Tangens: } \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

A négy kapu az alábbi funkciókat tölti be:

- **f:** felejtési kapu, egy a cella állapottal azonos méretű vektor, melynek minden eleme nulla és egy között van a szigmoid nemlinearitás hatására. A cella állapot vektorát ezzel a vektorral elemenként megszorozva a cella állapot bizonyos részleteit elfelejtjük.
- **g:** főkapu, amelynek feladata, hogy a bemenet aktuális értékéből és a cella kimenet korábbi értékéből kinyerje azokat a jellemzőket, amelyek a cella állapotban megjegyezhetők.
- **i:** bementi kapu, vektorával a főkapu vektorát elemenként szorozzuk, ezáltal kiválasztva a megjegyezhető jellemzőkből a releváns részeket, majd ezt a cella állapothoz hozzáadjuk.
- **o:** kimeneti kapu, megszűri a bemenet értékeit, csak a lényegeseket megtartva, majd veszi az elemenkénti szorzatát az aktuális cella állapot hiperbolikus tangens függvényével, így a cella aktuális kimenetét adva. (Szemenyei, 2021)

Az LSTM cella működése formalizálva:

$$i = \sigma \left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} + b_i \right)$$

$$f = \sigma \left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} + b_f \right)$$

$$o = \sigma \left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} + b_o \right)$$

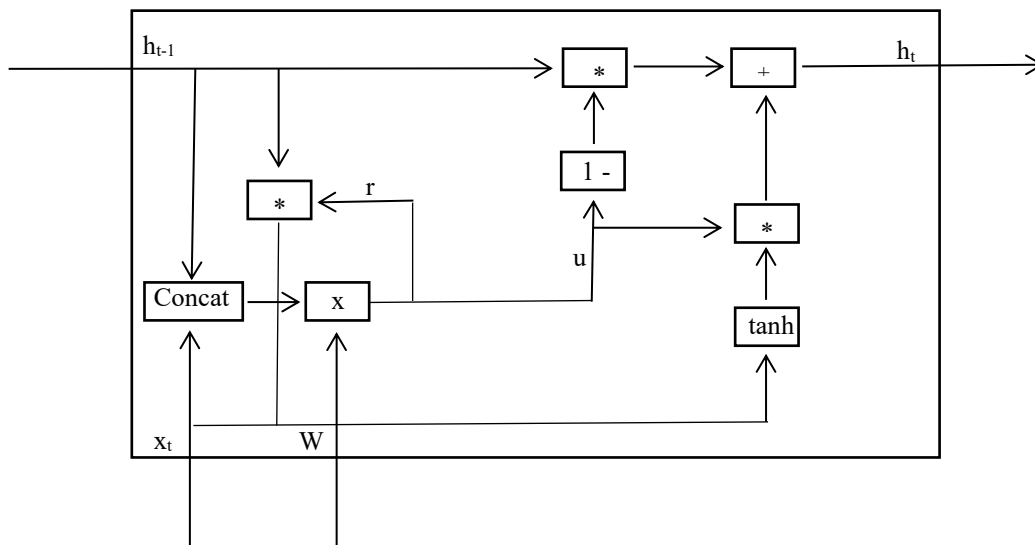
$$g = \tanh \left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} + b_g \right)$$

$$c_t = f * c_{t-1} + i * g$$

$$h_t = o * \tanh(c_t)$$

4.7 Kapuzott visszacsatolt cella

A kapuzott visszacsatolt cella (GRU – Gated Recurrent Unit) az LSTM cellához hasonló ötlet mentén lett megalkotva, azonban ez egy lényegesen egyszerűbb struktúra, amely a gyakorlati alkalmazás szempontjából előnyösebb, hiszen a kevesebb műveletet kell elvégezni az tanítás során, így a futási időt lényegesen tudja rövidíteni.



5. ábra: GRU cella struktúrája, (Olah, 2015)

A GRU cellában mindössze két kapu van:

- r: visszaállítási kapu. Azt határozza meg, hogy a múltbéli információkból mennyit felejtsünk el.
- u: frissítési kapu. Az LSTM hálózat felejtési és bementi kapujaihoz hasonlóan eldönti, hogy mely információt érdemes megtartani vagy elhagyni. (Cho, és mtsai., 2014)

Ebben az esetben is kizárólag a szigmoid és a hiperbolikus tangens aktivációs függvényt alkalmazzuk. Az GRU cella működése formalizálva (Olah, 2015):

$$u = \sigma\left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} + b_u\right)$$

$$r = \sigma\left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} + b_r\right)$$

$$h_t = (1 - u_t) * h_{t-1} + u_t * \tanh\left(W \begin{pmatrix} r_t * h_{t-1} \\ x_t \end{pmatrix}\right)$$

5 Adatok bemutatása

Dolgozatomban a modellezéshez szükséges magyar halandósági adatokat a Human Mortality Database (HMD, 2024) nyilvános oldaláról szereztem be. A megfigyelési időszaknak, az elérhetőség szerinti leghosszabb időszakot, 1950-2020-as időszakot választottam. Minden naptári évre minden életkor (0-tól 100-ig) mortalitási rátájának előállításához beszereztem az adott időszak során az adott populációban bekövetkezett összes (női és férfi adatok együttesen) halálozások számát, illetve a vizsgált időszak alatti központi kitettségeket.

A 3. fejezetben ismertetett logaritmált halandósági rátákat a beszerzett adatokkal előállítottam. A neurális hálózatok kimentének – eredményváltozójának – éppen a logaritmált halandósági rátát választottam pontosan úgy, mint a Lee-Carter modellnél. Bemenetének – magyarázóváltozóinak – pedig a logaritmált halandósági ráták késleltetett értékeit választottam. A maximális késleltetés 10 év volt. Egy megfigyelésre a becsléshez felhasznált változók az alábbi módon írhatók fel:

$$\text{Bemenet: } (\ln(m_{x,t-1}), \ln(m_{x,t-2}), \dots, \ln(m_{x,t-10}))$$

$$\text{Kimenet: } \ln(m_{x,t})$$

A Lee-Carter és a neurális hálós modellek teljesítményének összemérésére szükséges volt a rendelkezésre álló adatot a naptári évek szerint tanuló és tesztelő időszakokra szétbontani. A tanuló időszakon illesztettem – mind a Lee-Carter, mind a neurális hálózat – modelleket és a tesztelő időszakra előrejelzést készítettem az illesztett modellek alapján. A tanuló időszakot 1950 – 2005-nek, a tesztelő időszakot 2006 – 2020-nak választottam meg.

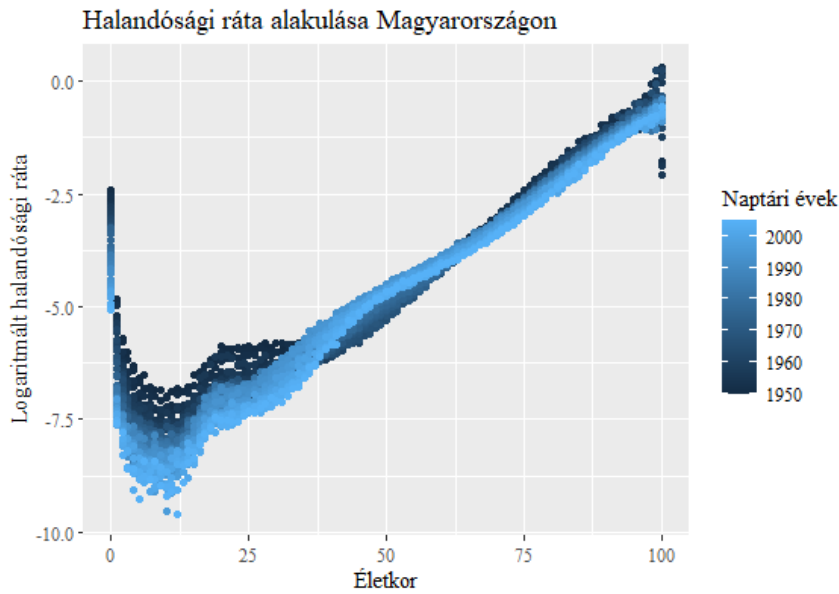
A Lee-Carter modell alkalmazásához további adatelőkészítésre nem volt szükség. A Lee-Carter modellt az R programozási nyelv demography (Hyndman, 2023) csomagjának segítségével állítottam elő.

Azonban a neurális hálózatok esetén szükséges volt a változókat lineáris transzformáció segítségével a $[0,1]$ intervallumra transzformálni, hogy ne egyetlen, nagyságrendben a legnagyobb változó dominálja a tanulási folyamatot. Továbbá a tanuló adatbázis 20%-a leválasztásra került validációs célokra. A validáló adatbázis megfigyelései véletlenszerűen kerültek kiválasztásra. A neurális hálózatok modellezésére a Python programozási nyelvet és a keras (Chollet, 2015) könyvtárát használtam, melyet nem csak lokálisan, de „felhőben” is futtattam. Utóbbi főként a hosszabb optimalizációs folyamatoknál alkalmaztam. Erre lehetőséget a Google Colab internetes GPU-hozzáférése nyújtott.

6 A Lee-Carter modell eredményeinek ismertetése

Ebben a fejezetben a Lee-Carter modell eredményeit kívánom áttekinteni. A modellhez az 1950 – 2005 közötti halandósági adatokat használtam fel és a 3. fejezetben leírt eljárással becsültem a teljes (női és férfi adatok együttesen) logaritmált halandósági rátát.

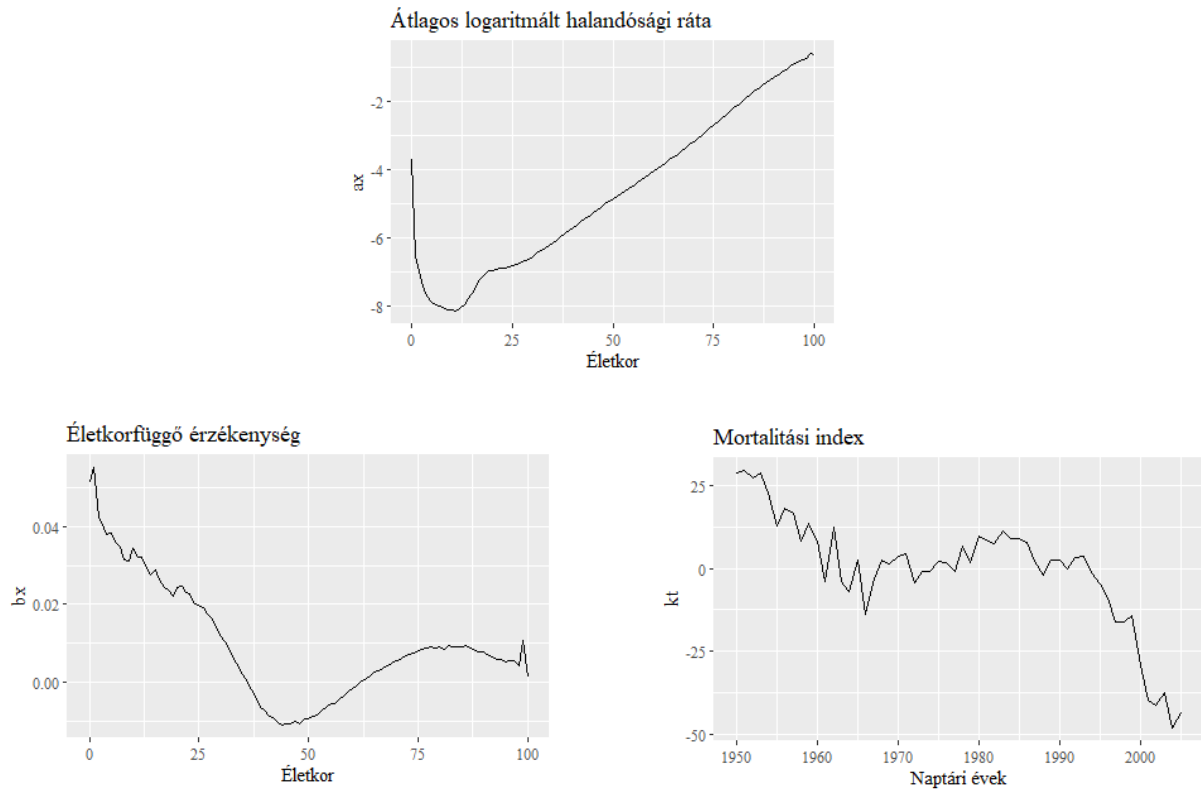
Tekintsük előbb a valós logaritmált halandósági rátákat életkoronként és naptári évenként.



6. ábra: Valós logaritmált halandósági ráta életkoronként és naptári évenként

Általánosan megállapítható, hogy átlagos szintje életkoronként más és más. Hosszú távon szinte minden életkorban csökkent értéke az elmúlt évek során, de ezen csökkenés üteme életkoronként eltérő. A Lee–Carter modell éppen ezen jelenség-hármaszt próbálja meg leírni. (Vékás, Halandósági táblák statisztikája, 2023)

A becsült paraméterek értékét életkorok vagy éppen naptári évek szerint a 7. ábra mutatja. Az átlagos logaritmált halandósági ráta a vártnak megfelelően alakult. Az életkorfüggő érzékenység láthatóan nagyon változó. Gyermekkorban a legmagasabb az értéke. Abszolút minimumát a középkorúak esetén veszi fel, majd újabb, de kisebb mértékű növekedést mutat az idősök körében. A mortalitási index mutat azonban a megszokottól leginkább eltérő viselkedést. Várakozásaim alapján a teljes időszakra csökkenő trendet követő értékeket kellene felvennie. Ez megfigyelhető 1950 – 1970 közötti időszakban, azonban 1970 – 1990 között stagnálást, illetve enyhe növekedést figyelhetünk meg. Az ezt követő időszakban újra csökkenő trendet követ az idősor.



7. ábra: Lee-Carter modell paraméterei

A modell egyetlen időfüggő paramétere a mortalitási index, ennek előrejelzése segítségével lehet előállítani a logaritmált halandósági ráta előrejelzését is. A Lee-Carter modell 2006 – 2020 időszaki előrejelzései a neurális hálózatok eredményeivel együtt kerülnek bemutatásra a A modellek eredményeinek összevetése⁸. fejezetben.

7 A neurális hálózatok eredményeinek ismertetése

7.1 Neurális hálózatok hiperparaméter-optimalizálása

A hiperparaméterek a modell olyan paraméterei, amelyek nem automatizált módon kerülnek meghatározásra, hanem magunk választhatjuk meg értéküket. Ezekkel finomhangolni lehet a rendszert az adott problémára specifikusan. (Havlik, 2016) A hiperparaméterek megfelelő megválasztása a mesterséges neurális hálózatoknál kiemelt jelentőségű. A jól megválasztott hiperparaméterek pontosabb becslésekhez vezetnek, ezzel csökkentve a hibát.

Minden neurális hálós architektúra esetén három rétegű hálót készítettem. Mindegyik egy bementi, egy rejtett és egy kimeneti rétegből állt. A kimeneti réteg aktivációs függvénye minden esetben a lineáris volt.

Továbbá minden modellbe beleépítettem egy korai megállás (early stopping) funkciót is, amely azt biztosította, hogy ha a tanulás során, az előre megadott lépésszám után sem csökken a hiba, akkor a tanulás leáll. Jogosan merül fel a kérdés, hogy ez miért szükséges. A tanulási folyamat alapvetően is sok időt vehet igénybe, azonban a hiperparaméter-optimalizációs folyamat különösen időigényes. Mivel ezek tetszőlegesen meghatározott értékek, így megválasztásuk próbálgatással lehetséges. Ha több különböző hiperparaméterünk van és ezekre több lehetséges értéket is szeretnénk megvizsgálni, akkor a modellt annyiszor fogjuk újratanítani ahány lehetséges kombinációt elő tudunk állítani. Ez pedig a modellek komplexitásától függően több órától kezdve akár napokig vagy hetekig is eltarthat.

A hiperparaméterek és azok lehetséges értékei, amelyek a dolgozat során optimalizálásra kerültek az alábbiak.

Rejtett réteg neuronjainak száma: $n = (1, 2, 3, 4, 5)$

Tanulási ráta: $\lambda = (0,01, 0,005, 0,001)$

Minikötegek nagysága: $b = (16, 32, 64, 128)$

A korai megállás 75 lépésszám után következhetett be, amennyiben az egymást követő 75 lépésben nem csökkent a hiba. A maximális lépésszám minden modell esetén 2000 volt.

Optimálisnak akkor tekintettem egy hiperparaméter-kombinációt, ha az eredményezte a validáló adatokon elért legkisebb MSE-t. Ezek alapján az egyes modellekre az alábbi táblázat tartalmazza az optimális hiperparaméter választást.

Modell	Rejtett réteg neuronjainak száma (n)	Tanulási ráta (λ)	Minikötegek nagysága (b)
Feed-Forward	3	0,001	16
RNN	3	0,001	64
LSTM	5	0,001	64
GRU	3	0,005	128

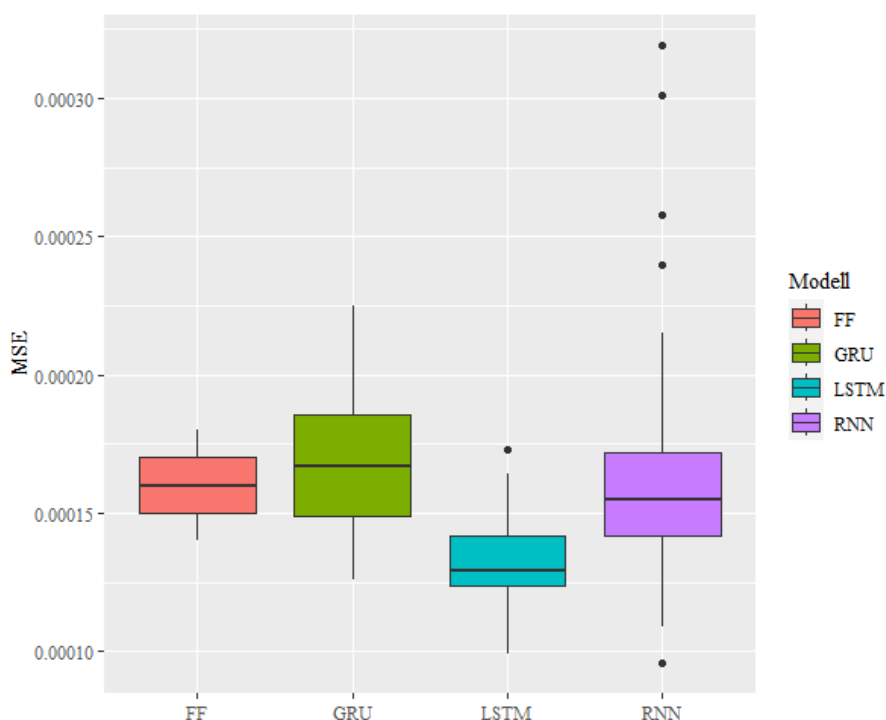
2. táblázat: Optimális hiperparaméterek

A kapott eredmények alapján elmondható, hogy a modellek többsége 3 neuront tartalmaz a rejtett rétegben, ez alól csak az LSTM architektúrájú modell képez kivételt. A tanulási ráta is csak a GRU architektúrájú modell esetén különbözik 0,001-től. A minikötegek nagysága nagyon különbözően alakult.

A fenti hiperparaméter megválasztású és struktúrájú modellekkel dolgoztam a továbbiakban. Ezen modellek megkonstruálása után a paraméterek száma a Feed-Forward modell esetén 37, RNN esetén 19, LSTM esetén 146, GRU esetén 58 volt. A rendelkezésre álló megfigyelések száma elegendő volt bármely imént említett paraméterszám mellett.

7.2 A hiba nagyságának alakulása

A neurális hálózatok tanulási folyamata során a háló kezdeti súlyai mindig véletlenszerűen kerülnek megválasztásra. Éppen emiatt a modell újratanításonként különböző optimális súlyokat és MSE-t adhat eredményül. A tesztelő halmazon elért hibát, MSE-t, mindig a (nem logaritmált) halandósági rátákra állapítottam meg.



8. ábra: MSE értékek alakulása modelltípusonként

A 8. ábra szemlélteti a tesztelő halmazon elért MSE alakulását. A legnagyobb és a legkisebb MSE-t is az RNN típusú modell eredményezte, tehát szóródása is ennek volt a legnagyobb. A legkisebb medián értéket az LSTM típusú modellek vették fel, várakozásaimnak megfelelően, hiszen ez a struktúra tipikusan az idősoros adatok feldolgozására lett megalkotva. Ami meglepőnek bizonyult, hogy a GRU típusú neurális hálózatok nem teljesítettek jobban a szimpla Feed-Forward modelleknél.

Típusonként a 10 legkisebb hibájú modell átlagát az 3. táblázatban foglaltam össze.

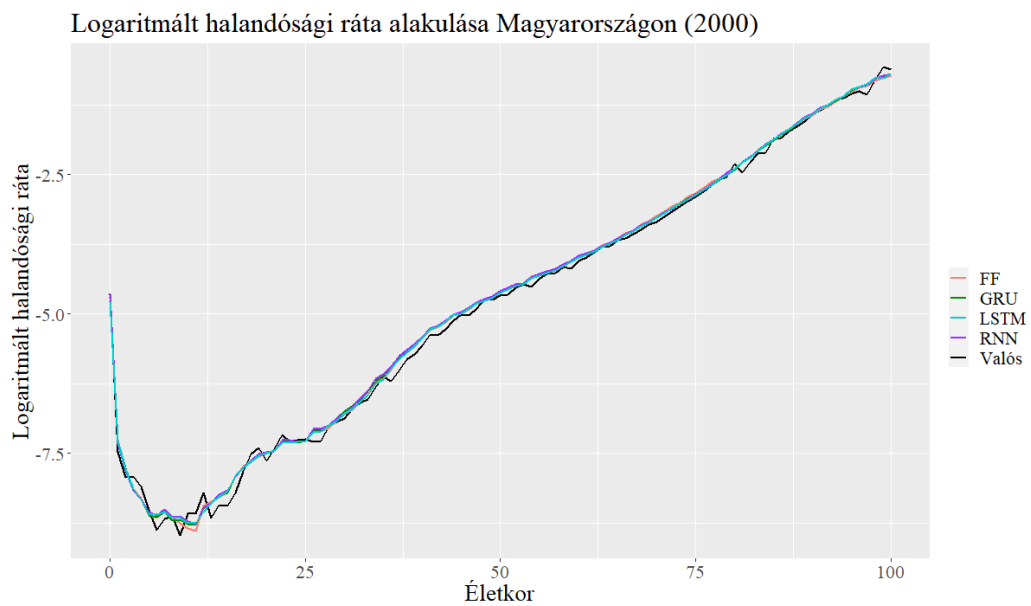
Modell	Feed-Forward	RNN	LSTM	GRU
MSE	1,48e-4	1,09e-4	1,11e-4	1,39e-4

3. táblázat: 10 legkisebb MSE-jű modell átlaga típusonként

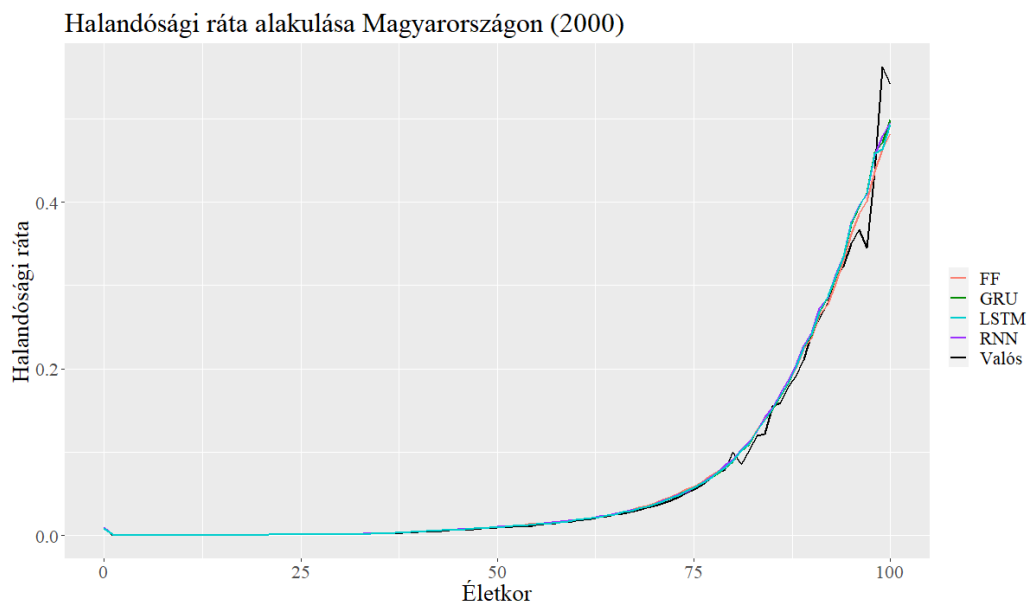
Annak ellenére, hogy az RNN típusú modellek átlagban alacsonyabb MSE-t eredményeztek, az LSTM típusú modellek hibájának szóródása lényegesen alacsonyabb volt és stabilan kisebb hibával becsült. Az LSTM esetén a medián 0,000129, míg RNN esetén ennél magasabb 0,000155 volt. Mivel a medián egy robosztus mutató, így erre hivatkozva – és tekintve, hogy minimális különbség volt az átlagos MSE-k között is – az LSTM modellt használtam az előrejelzések készítéséhez.

7.3 Becsült mortalitási ráták alakulása

Tekintetem a tanuló adatbázis becsült értékeit és összehasonlítottam a valósakkal. Ezt keresztmetszetben, naptári évenként vizsgáltam meg. Az eltérések egyik naptári év esetén sem voltak szignifikánsan különbözőek a többitől. A 9. ábra– 10. ábraán a 2000-es naptári évhez tartozó logaritmált és transzformálatlan halandósági rátákat ábrázoltam életkoronként és modellenként.



9. ábra: Logaritmált halandósági ráta életkoronként (2000)



10. ábra: Halandósági ráta életkoronként (2000)

Mindegyik modell jól illeszkedik a fekete, valós értékekhez. Az eltérések minimálisak a modellek között.

8 A modellek eredményeinek összevetése

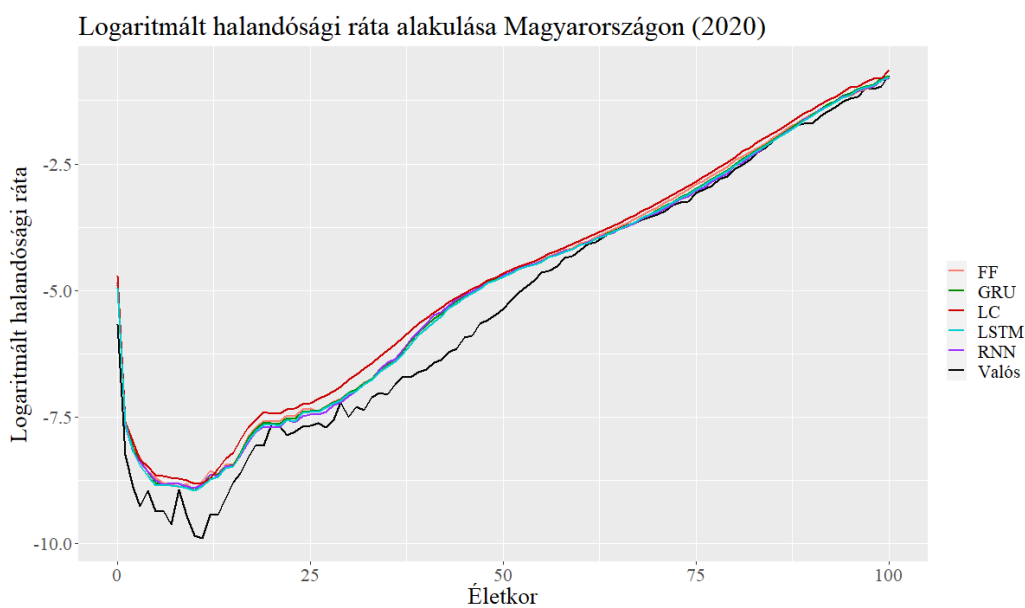
A Lee-Carter és neurális hálózatok, mint halandóság előrejelzési modellek kerülnek összevetésre ebben a fejezetben. Az összehasonlítás első szempontja a tesztelő halmazon elért predikciós hiba volt.

Modell	Feed-Forward	RNN	LSTM	GRU	Lee-Carter
MSE	1,48e-4	1,09e-4	1,11e-4	1,39e-4	6,57e-4

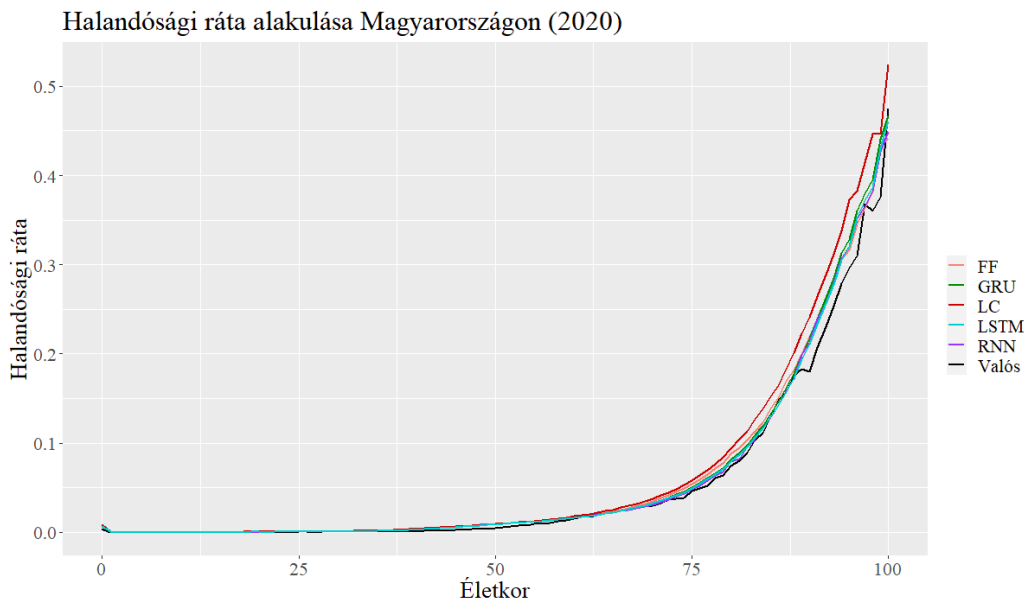
4. táblázat: MSE modellenként

A tesztelő halmazon megállapított MSE a Lee-Carter modell esetén volt a legmagasabb. Még a nem kifejezetten idősoros adatok feldolgozására alkalmas Feed-Forward neurális hálózat is kisebb hibájú predikciókat adott.

A tesztelő halmaz predikcióinak elkészítésénél fontos megjegyezni, hogy az előrejelzéseket mindig csak lépésről lépésre, a következő évre szabad elvégezni, ahogy azt írja (Richman & Wüthrich, 2019). A későbbi évek előrejelzéseit a korábbi évek becslései segítségével kell előállítani. A 2005-ös év mortalitási rátáit előre tudjuk jelezni a 2004 és 1994 közötti tényleges halandósági rátákkal. A 2006-os év előrejelzésénél azonban már nem használhatjuk a valós 2005-ös halandósági rátákat, hanem a 2005-ös előrejelzéseket kell használni. (Szentkereszti & Vékás, 2022) Ezeket figyelembe véve tekintettem – a neurális hálózatoknál alkalmazott módon – a 2020-as naptári évhez tartozó logaritmált és transzformálatlan halandósági rátát életkoronként és modellenként.



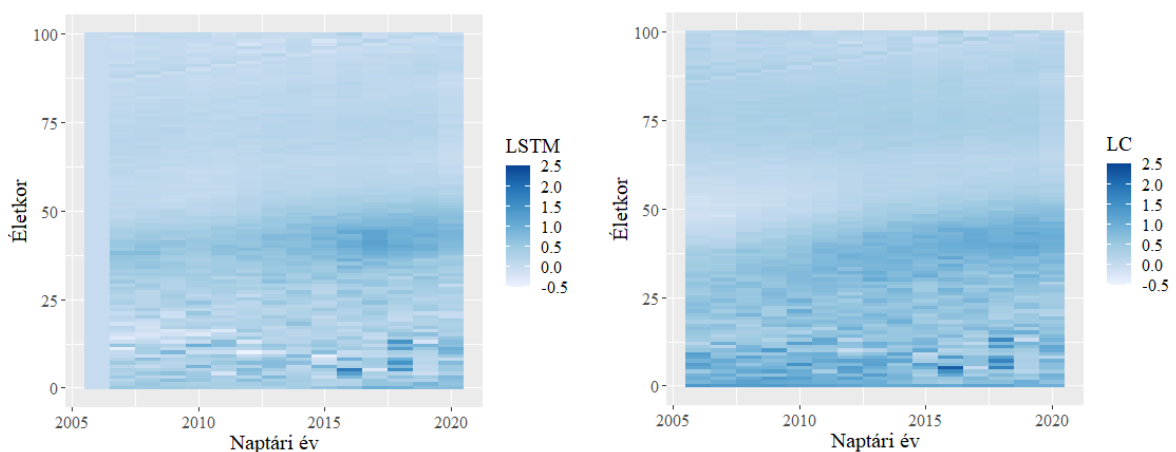
11. ábra: Logaritmált halandósági ráta (2020)



12. ábra: Halandósági ráta (2020)

Megfigyelhető, hogy míg a különböző struktúrájú neurális hálózatok egymáshoz nagyon hasonló pontossággal becsülnek, addig a Lee-Carter ezektől szignifikánsan nagyobb hibájú becslést ad. A logaritmált halandósági ráták esetén jobban kivehető, hogy a becslések a középkorúak esetén térnek el leginkább a valóstól.

Megvizsgáltam, hogy ezek a reziduálisok a többi életkor esetén hogyan alakultak a tesztelő időszakon. A neurális hálós modellek közül, csak az LSTM reziduálisait vizsgáltam tekintve, hogy a modellek pontossága csak minimálisan különbözött és későbbi számításaimhoz is ezt az architektúrát alkalmaztam.

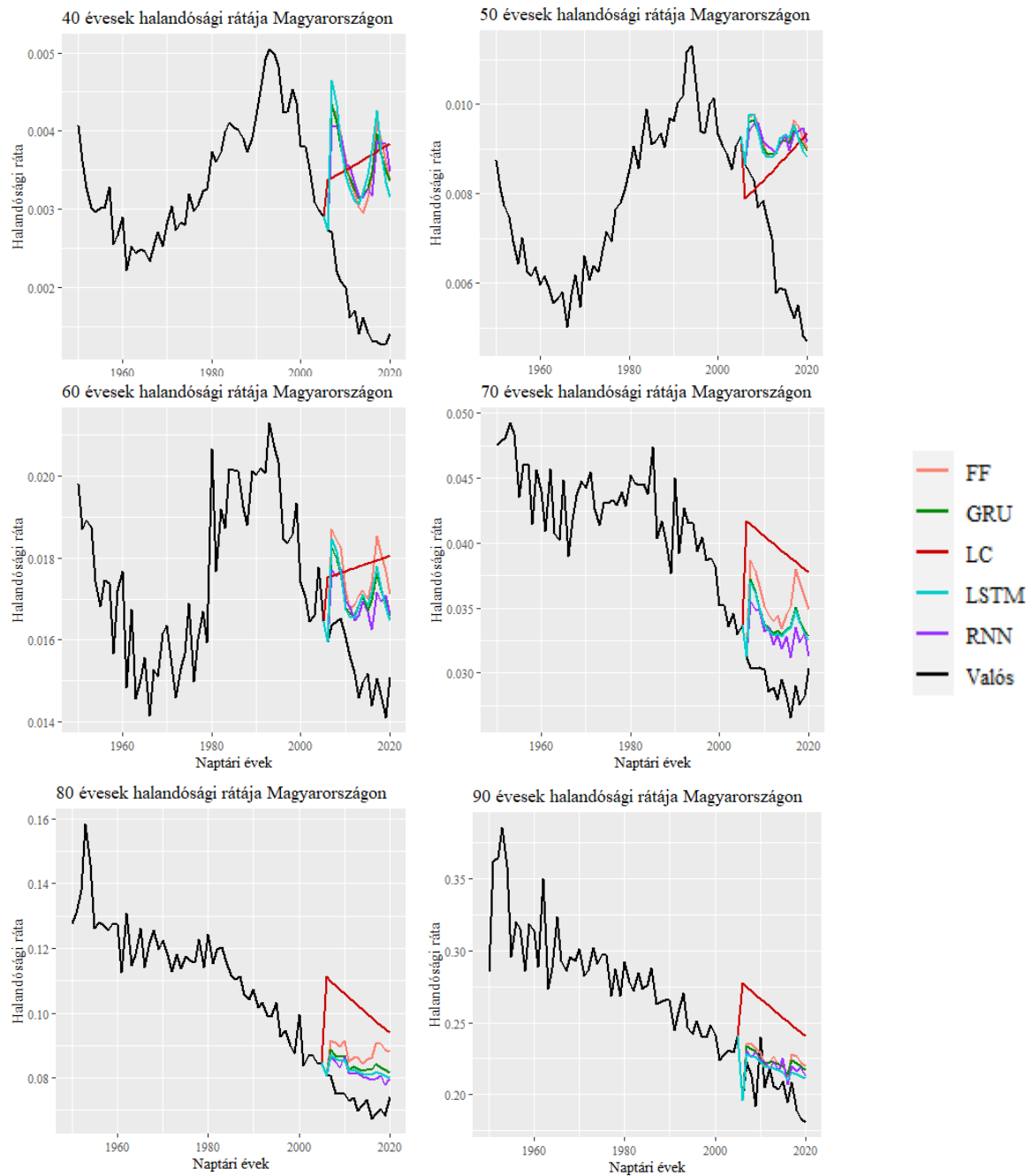


13. ábra: A Lee-Carter és az LSTM struktúrájú neurális hálós modell reziduálisai

A középkorúak reziduálisai mind a Lee – Carter modell, mind az LSTM architektúra esetén nagyobb eltérést mutattak a többi életkorhoz képest. Az LSTM esetén viszont valamennyivel

kisebbsnek bizonyultak ezek az eltérések. Éppen ez figyelhető meg a 11. ábraán. Illetve a fiatalabb korosztályok esetén is a reziduális ábra világosabb az LSTM alkalmazása esetén, így azt lehet mondani, hogy a fiatalabb korosztályokra az LSTM alacsonyabb hibával becsül. Az időskori logaritmált halandósági rátát közel egyformán alacsony hibával becsülték.

Jó összehasonlítási alap lehet még a tesztelő halmazra elkészített predikciókat nemcsak keresztmetszetben, de korosztályonként is elkészíteni.



14. ábra: Halandósági ráta életkoronként

A 40-90 éves korosztályokra (10 évenként) vizsgáltam meg a tesztelő minta becsléseit. A legnagyobb hiba – bármely modell esetén – valóban a 40-50-60-as életkorok esetén van, nem meglepő módon. Megfigyelhetően a neurális hálós modellek adnak pontosabb becsléseket, köztük is az LSTM és az RNN architektúrájúak a legkisebb hibájúak. Megállapítható, hogy míg a Lee-Carter legfeljebb az idősor trendjét tudja előrejelzésként adni, addig a neurális hálózatok képesek az egyes évek közti ingadozást is jól felvenni.

Az összehasonlítás alapján az mondható el, hogy a halandóságot a neurális hálózatok pontosabban becsülték ezen tanító-tesztelő adatok megválasztása esetén. Köztük is az LSTM és RNN hálózatok bizonyultak a legkisebb hibájúnak.

9 Előrejelzések jósága

9.1 Az élettartam-kockázatról

Az elemzés során megfigyelhető volt, hogy a mortalitási ráták átlagosan minden életkorban csökkenő trendet mutattak. E folyamat jelentős módszertani és pénzügyi nehézségek elé állítja a nyugdíj- és életbiztosítási területen működő intézményeket, melyek pénzáramlásainak előrejelzése során feltételezik, hogy – egy bizonyos naptári évre vonatkozó – koréves halálozási valószínűségek a jövőben változatlanok maradnak. Azonban a valóságban az életjáradékok tulajdonosai a halandóság csökkenése következtében a kalkulálnál nagyobb valószínűséggel érik meg a járadékfizetési időpontokat, ami jelentős, előre nem kalkulált többletkifizetést jelent a járadékszolgáltatóknak. Változatlan halandóságot feltételezve csak pontatlanul tudják előrejelezni jövőbeli bevételeiket, valamint kiadásait, ezzel komoly tervezési hibát elkövetve. Az ebből fakadó problémát hívjuk élettartam-kockázatnak. (Vékás, Nyugdíjcélú életjáradékok élettartam-kockázata az általánosított korcsoport-időszak-kohorsz modellkeretben, 2017)

Az élettartam-kockázat jelenségét számszerűsíteni lehet olyan módon, hogy dinamikus – kohorsz – szemléletű halandósági táblák segítségével kiszámoljuk a várható hátralévő élettartam és az egyszeri díjas életjáradék értékét és összehasonlítjuk a statikus – egy bizonyos naptári évre vonatkozó – halandósági tábla által számolt értékekkel.

Dinamikus szemléletű halandósági táblákat viszont csak előrejelzett adatokkal lehet elkészíteni bizonyos korosztályokra. Így a dolgozatban bemutatott legjobb halandóság előrejelző modell segítségével elkészítettem a 65 éves korosztályra vonatkozó kohorsz halandósági táblát. Az ebből számolt várható hátralévő élettartam és az egyszeri díjas életjáradék értékeit pedig összevettem a (Májér & Kovács, 2011) cikkében alkalmazott módosított Lee-Carter és a (Vékás, Nyugdíjcélú életjáradékok élettartam-kockázata az általánosított korcsoport-időszak-kohorsz modellkeretben, 2017) GAPC-modellkeret által számolt értékekkel.

Az összehasonlítás legfőbb célja az volt, hogy megvizsgáljam, hogy a neurális hálós modell hosszú távon is képes-e megbízható és jól értelmezhető becslést adni. Továbbá, hogy ez alkalmas-e az élettartam-kockázat kiszűrésére a megfelelő mutatók számításánál.

9.2 Várható hátralévő élettartam és életjáradék számítása

Az x éves korban várható hátralévő élettartamot megadó összefüggés statikus halandósági tábla alapján (Banyár, 2003):

$$e_x = \sum_{i=1}^{\omega-x} \prod_{j=0}^{i-1} (1 - q_{x+j}) + \frac{1}{2} \quad (x = 0, 1, \dots, \omega)$$

ahol q_x jelöli az x éves életkor szerinti halálozási valószínűséget és ω a feltételezett legmagasabb elérhető életkort. Esetemben ezt 100-nak tétéleztem fel. Ha dinamikus halandósági táblával írjuk fel, akkor az összefüggés az alábbiak szerint változik:

$$e_x = \sum_{i=1}^{\omega-x} \prod_{j=0}^{i-1} (1 - q_{x+j,T+j}) + \frac{1}{2} \quad (x = 0, 1, \dots, \omega)$$

ahol q_{xt} az x éves életkorhoz tartozó halálozási valószínűség a t . naptári évben. T az aktuális naptári évet jelöli.

Az x éves korú egyének azonnal induló, élethosszig tartó, évi egy forint összegű életjáradékának egyszeri nettó díjképlete az aktuáriusi ekvivalenciaelv szerint (Banyár, 2003):

$$\ddot{a}_x = \sum_{i=0}^{\omega-x} \left[v^i \prod_{j=0}^{i-1} (1 - q_{x+j}) \right] \quad (x = 0, 1, \dots, \omega)$$

ahol v a technikai kamatláb segítségével előállított diszkontfaktor. A nettó díj azt az életjáradékért cserébe nyújtandó egyösszegű befizetést adja meg, amely mellett a járadékszolgáltató a technikai kamatlábnak megfelelő, rögzített éves befektetési hozam feltételezése mellett, a járadékfizetésen kívüli egyéb költségek figyelembevétele nélkül várhatóan nulla profitot realizál. (Vékás, Nyugdíjcélú életjáradékok élettartam-kockázata az általánosított korcsoport-időszak-kohorsz modellkeretben, 2017)

Ugyanez kohorsz halandósági ráták használatával:

$$\ddot{a}_x = \sum_{i=0}^{\omega-x} \left[v^i \prod_{j=0}^{i-1} (1 - q_{x+j,T+j}) \right] \quad (x = 0, 1, \dots, \omega)$$

A q_x , azaz x éves életkorhoz tartozó halálozási valószínűségek maximum likelihood becslése, az általam alkalmazott központi halálozási rátákból, az alábbi formulával kapható meg (Vékás, Halandósági táblák statisztikája, 2023):

$$\hat{q}_x = 1 - e^{-m_x}.$$

9.3 Az életjáradékokkal kapcsolatos eredmények

Az élettartam kockázat számszerűsítésére az LSTM struktúrájú neurális hálózatot a teljes megfigyelési időszakra (1960 – 2020) újratanítottam és így készítettem predikciókat egészen 2060-ig. Ezen becslésekből pedig a 9.2 fejezetben leírt módon elkészítettem a 65 évesek kohorsz szemléletű halandósági tábláját, amiből aztán meghatároztam az e_{65} – 65 éves korban várható hátralévő élettartam – és \ddot{a}_{65} – 65 éves korú egyének azonnal induló, élethosszig tartó, évi egy forint összegű életjáradékának egyszeri nettó díja – értékét.

A technikai kamatlábat a Magyar Nemzeti Bank vonatkozó rendelete (Magyar Közlöny, 2022) alapján a maximális 4%-nak választottam meg. Vékás cikkében 2,3%, Májér-Kovács pedig 3%-os technikai kamatláb feltételezéssel éltek.

A becslésekhez konfidencia intervallumot is meghatároztam. Az intervallum széleit következőképpen választottam meg:

$$CI = \bar{q}_{x,t} \pm \frac{1}{2}SD(\hat{q}_{x,t})$$

azaz az x éves életkor t . évhez tartozó halálozási valószínűség becslések átlaga – éppen azok az értékek, amiket predikciók elkészítéséhez is alkalmaztam – megnövelve vagy éppen csökkentve ugyanezen becslések szórásának felével.

Számításba akartam venni a koronavírus-járvány halandóságra gyakorolt hatását, így a predikciókat két bázisidőszakra is elkészítettem: 1950–2019 és a 1950–2020 közötti időszakra. Ennek célja mindössze az volt, hogy a predikciók az esetleges torzítástól mentesen is előálljanak.

Hogy az összehasonlítást megfelelően el tudjam végezni, meghatároztam a 2019-es és 2020-as naptári évre vonatkozó statikus halandósági táblák alapján is az e_{65} és \ddot{a}_{65} értékeket. Az összehasonlításra alapvetően a 1950–2020-as bázisidőszak szerinti eredményeket használtam. Továbbá tekintetem a Vékás és Májér-Kovács szerzők által meghatározott értékeket.

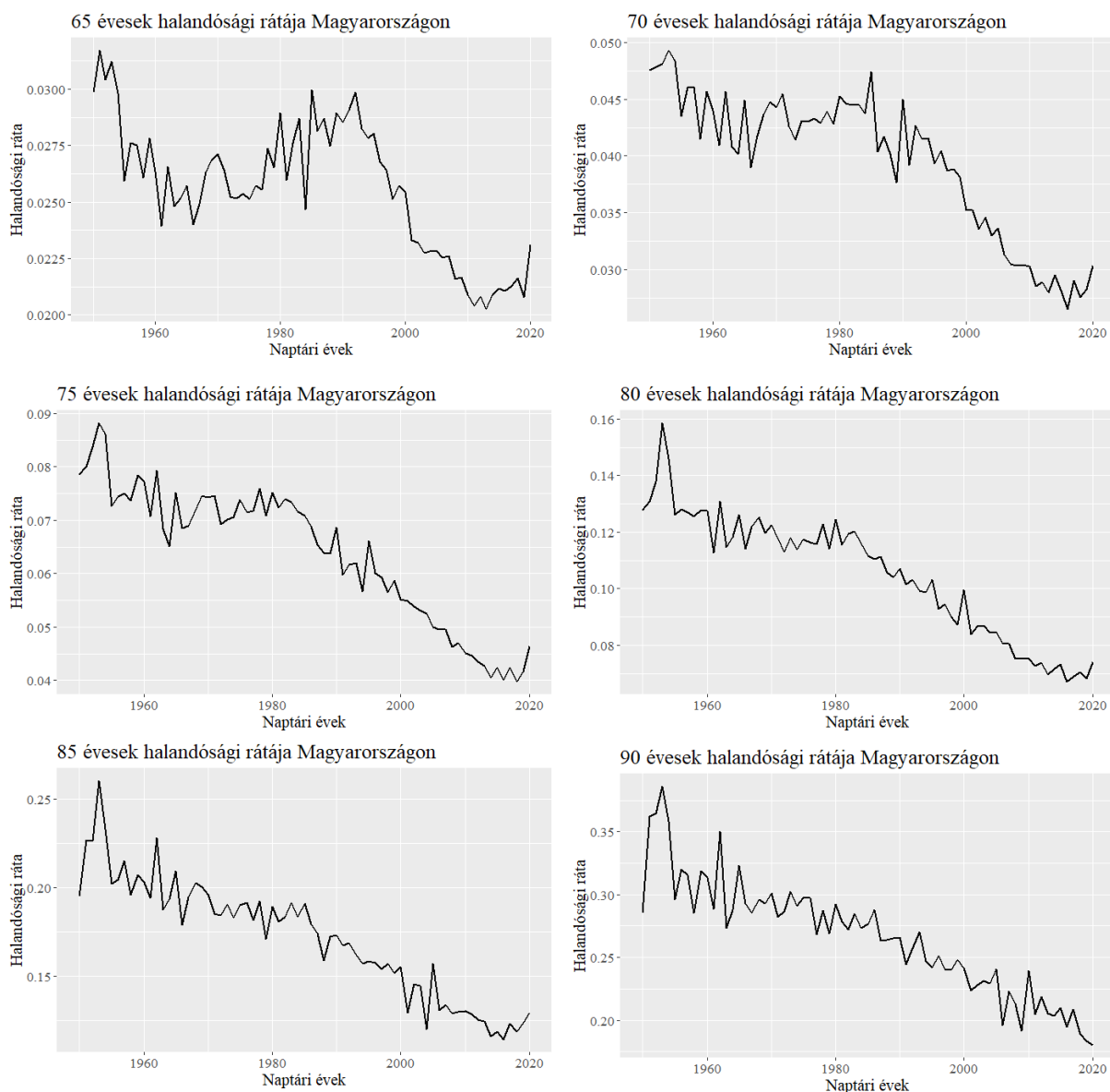
Mennyiség	Keresztmetszeti érték	Várható érték (konfidenciaintervallum)	Statikus hiba (százalék)
	Saját számítás (bázisidőszak: 1950–2020)		
e_{65} (év)	16,19	15,41 (15,33;15,48)	5,06
\ddot{a}_{65} (forint)	11,73	11,33 (11,29;11,36)	3,53

Saját számítás (bázisidőszak: 1950–2019)			
e_{65} (év)	16,85	15,41 (15,37;15,50)	9,34
\ddot{a}_{65} (forint)	12,07	11,33 (11,31;11,38)	6,51
Vékás (2017) (bázisidőszak: 1975–2014)			
e_{65} (év)	16,47	18,21 (16,61; 19,80)	-9,51
\ddot{a}_{65} (forint)	13,72	14,78 (13,83;15,72)	-6,43
Májjer-Kovács (2011) (bázisidőszak: 1970–2006)			
e_{65} (év)	15,39	16,43 (15,12; 17,83)	-6,33
\ddot{a}_{65} (forint)	11,87	12,43 (11,70; 13,17)	-4,50

5. táblázat: Összehasonlítás: a 65 éves korban várható hátralevő élettartam és az életjáradék egyszeri nettó díja (Májjer & Kovács, 2011), (Vékás, Nyugdíjcélú életjáradékok élettartam-kockázata az általánosított korcsoport-időszak-kohorsz modellkeretben, 2017)

Ami egyértelműen megállapítható, hogy a 2014-es keresztmetszeti szemléletű halandósági táblához képest a 2020-as szerinti a 65 éves korban várható hátralevő élettartam 0,28 évvel és az életjáradék egyszeri nettó díja 1,99 forinttal csökkent. Utóbbi köszönhető a hátralevő várható élettartam csökkenésének, illetve a technikai kamatláb növelésének. Ezen eredmények már alapvetően is meglepőek voltak, mindkét esetben növekedést vártam volna el keresztmetszetben. Számításaim szerinti e_{65} és \ddot{a}_{65} értékeket nemcsak a Vékás és Májjer-Kovács cikkekben közölt kohorsz szemléletű halandósági előrejelzések alapján számolt értékek, de a periódus tábla általi becslések is meghaladják. Ezt azt jelenti, hogy a neurális hálózat kisebb mértékű növekedést is vár a 65 évesek halandóságára, ami magyarázza az alacsonyabb nettó díját.

Az halandósági ráták ábráit jobban megvizsgálva arra a következtetésre jutottam, hogy valójában ezek nem irreális eredmények. Már a koronavírus-járvány időszakát megelőző 10 évben is nőtt vagy stagnált a halandóság a legtöbb 65 éves kor feletti korosztályban.



15. ábra: Halandósági ráta alakulása a 65 éves életkor feletti korosztályokban

Erre magyarázatul szolgálhat Omran elmélete az epidemológiai átmenetekről. (Omran, A theory of the epidemiology of Population change, 1971) Szerinte minden társadalom a modernizációs folyamat három korszakán megy keresztül. Az első szakasz a 18. század közepéig tartó premodern időszak a pestis és éhezés kora. A második szakaszt a visszahúzódó járványok korszakaként írta le. A harmadik korszakot a degeneratív és ember által okozott betegségek dominanciája jellemzi. A korszakok egészségjavulását eltérő tényezők befolyásolják. A harmadik szakaszt illetően az orvosi és közegészségügyi intézkedések, a megelőző és gyógyító eljárások, a jobb közegészségügyi intézkedések, a védőoltások (megelőzés) és az oki/célzott terápiák elterjedését emelte ki. Közel harminc évvel az első közleménye megjelenése után elmélete revideálására vállalkozott, ennek részeként egy új, negyedik epidemológiai korszak, „A csökkenő kardiovaszkuláris mortalitás idősödés, életmód-változás,

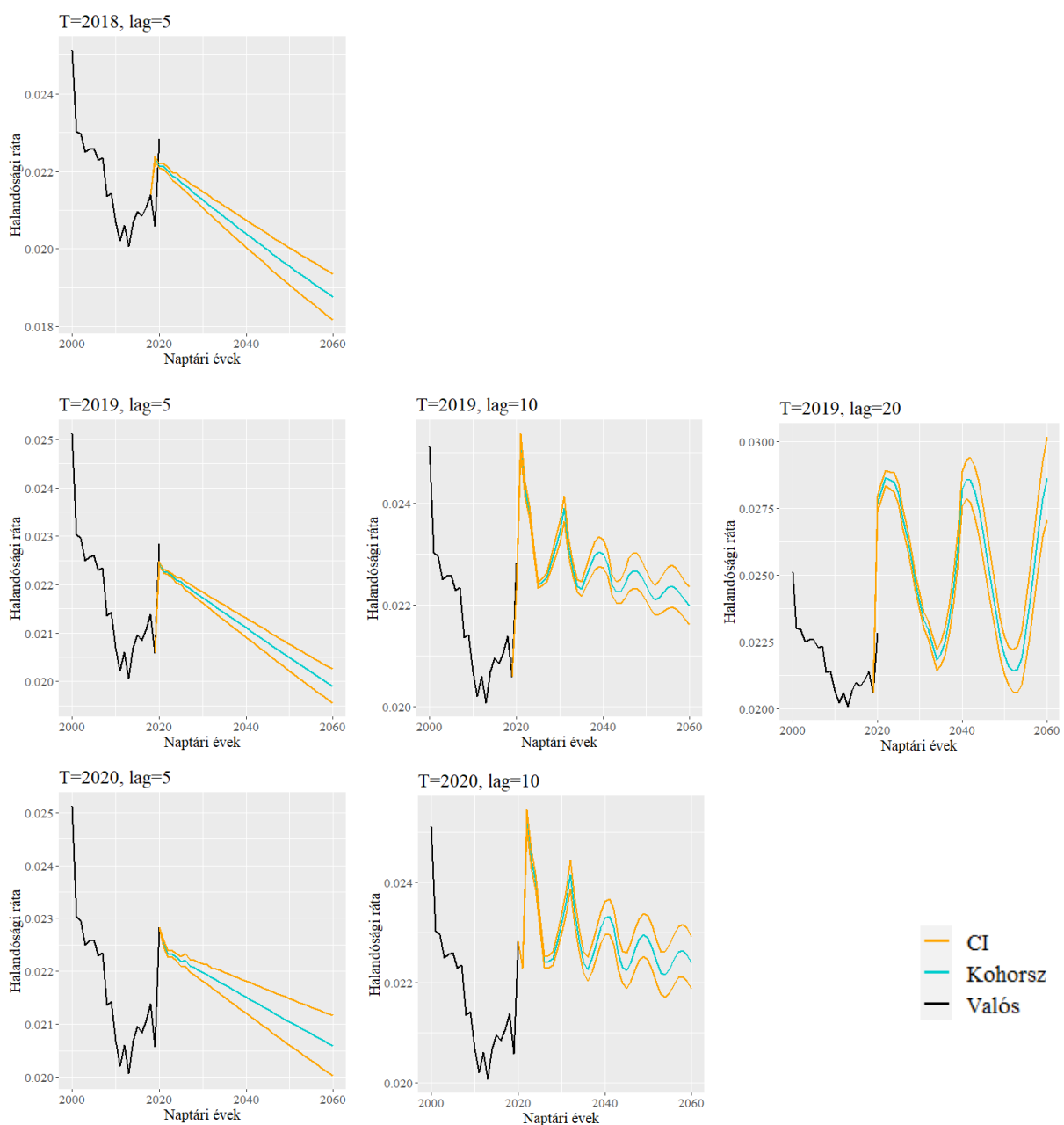
kialakuló és újjászülető betegségek” bevezetését javasolta. (Omran, The epidemiologic transition theory. A Preliminary update., 1983) A klasszikus epidemiológiai átmenet elmélete számára komoly kihívást jelentett, hogy a szocialista országok várható élettartamai folyamatosan távolodtak a nyugat-európai trendektől. A közép- és kelet-európai országok megfeneklettek az epidemiológiai fejlődés harmadik szakaszában, képtelenek voltak átlépni a következő fázisba. A keringési rendszer okozta halandóság a legmeghatározóbb, de nem az egyetlen epidemiológiai magyarázata a térség leszakadásának. (Caselli, Meslé, & Vallin, 2002) Az egészségkárosító magatartási formák széles körben elfogadottak voltak, és a mindennapi élet részévé váltak. (Cockerham, 1999) Az egészségmegőrzés feltételei korlátozottak voltak az államszocialista korszakban. (Józan, 2003a) A központosított társadalmi rendszer hátrányosan érintette az egyéneket abban, hogy egészségmagatartásuk, életmódjuk változásán keresztül felelősséget vállaljanak saját egészségükért. (Bálint, Mennyire illeszkedik a magyar halandóság alakulása az epidemiológiai átmenet elméleteihez?, 2016)

Azt az elméletet, hogy Magyarország is ebben a harmadik fázisban ragadt alátámasztja Bálint és Kovács tanulmánya. (Bálint & Kovács, Halandóság, 2018) A korszpecifikus halálozásokat tekintve a szerzőpáros is arra megállapításra jutott, hogy a 60 év feletti korosztályok halandósága stagnál, estenként nő a cikk megírását megelőző években. Arról is írnak, hogy ezekben a korosztályokban a megfigyelt időszakban a legtöbb környező országhoz és az Európai Unióhoz hasonlítva a magyar halálozások száma magasabb volt. Az okspecifikus halálozás elemzésénél pedig kiemelik, hogy Magyarországon – hasonlóan a legtöbb európai országhoz – a keringési rendszer betegségeiben halnak meg a legtöbben. Továbbá a jobb életkilátásokkal rendelkező országok között Magyarországon a második legnagyobb a daganatos betegségből eredő halálozás.

Többek között a fentiekben leírt demográfiai megfontolások indokolhatják a halandóság stagnálást a 65 év feletti korosztályokban.

A 2019-es periódus halandósági tábla alapján, az 65 éves korban várható hátralévő élettartam a 2014-eshez képest növekedést, az életjáradékának egyszeri nettó díja csökkenést mutat. Utóbbi a technikai kamatláb jelentős mértékű (1,7 százalékpontos) növekedésének köszönhető. 2,3%-os technikai kamatláb feltételezés mellett a 2019-es periódus tábla által megállapított halandósági valószínűségekkel számolt járadék értéke 13.93 Ft lenne. Ez pedig már nagyobb, mint a 2014-ben megállapított járadék értéke.

A kohorsz szemléletű halandósági tábla ebben az esetben is alacsonyabb várható hátralévő élettartamot és nettó díjat jósolt, mint a 2019-es periódus tábla szerinti értékek. Tehát hiába szűrtem ki a koronavírus járványból adódó többlethalandóságot, még mindig magasabb halandósági rátát jósol a modell, mint a keresztmetszeti eredmények. Ez annak köszönhető, hogy az elmúlt 10 időszak alapján készít predikciót a modell, amiben enyhe növekedést és stagnálást állapít meg. Emiatt megnéztem különböző késleltetésű bementi változókat alkalmazó modellek előrejelzéseit is. Megvizsgáltam, hogy a hosszabb vagy rövidebb késleltetés ad-e jobb előrejelzést. Továbbá megfigyeltem, hogy az előrejelzés kezdeti időpontjára nagyon érzékeny a neurális hálózat.



16. ábra: 65 évesek halálozási valószínűségeinek előrejelzése

A késleltetés növelésével, csak az előrejelzések szóródása lett nagyobb, míg az alacsonyabb késleltetés már a várt, trendszerű csökkenést mutatta. Azonban ebben az esetben sem volt mindegy, hogy honnan indul a predikció. A historikus minimum alá csak 2018-ig megfigyelt adatok esetén esett az előrejelzés értéke, azonban ebben az esetben is a kohorsz szemléletű tábla alacsonyabb értékeket mutatott, mint az adott évi periódus.

10 Összefoglalás

Dolgozatom célja az volt, hogy találjak olyan gépi tanulási modellt, amely az előrejelzések pontosságában el tud érni olyan jó eredményeket, mint a Lee-Carter halandóság-előrejelző modell. Ezt sikerült is belátnom, hiszen a neurális hálózat alapú modellek rendre kisebb hibával jelezték előre a korszecifikus halandósági rátákat a megfigyelt időszakon. Köztük is az LSTM és RNN hálózatok bizonyultak a legkisebb hibájúnak. A hosszútávú előrejelzésekhez az LSTM architektúrájú hálózatot használtam.

Az élettartam-kockázat számszerűsítése során a dinamikus halandósági táblákból előállított várható hátralévő élettartam és egyszeri nettó díj értékére a periódus táblából számolt értékeknél alacsonyabb számokat kaptam. Tehát a dinamikus előrejelzések szerint a halandóság magasabb lesz, mint az adott évi periódus tábla szerinti halandóság. Mivel ezek az eredmények a várakozásaimmal ellentétesen alakultak, így jobban szemügyre vettem a valós mortalitási rátákat. Ekkor figyeltem fel a halandósági ráta stagnálására, enyhe növekedésére az elmúlt 10 évben a 65 év feletti korosztályokban. Ennek egy lehetséges magyarázatára Omran tanulmányában leltem rá. Eszerint a posztszovjet országokban megfigyelhető, hogy a degeneratív és ember által okozott betegségekből eredő halálozások korszakából nem tudnak átlépni a negyedik korszakba, amely éppen ezek visszaszorulásából adódó további halandóság csökkenést jelentene. Hangsúlyozva, hogy ez éppen az életmódváltással, az egészségkárosító magatartások elhagyásával és prevencióval érhető el.

A neurális hálózat alapvetően kisebb hibával tudott előrejelzést készíteni az ismert időszakokra, azonban hosszú távú előrejelzése bizonytalan és megkérdőjelezhető becsléseket ad. Ennek ellenére a neurális hálózatok mellett szólhat, hogy a Lee-Carter modell nagyon korlátos abban a tekintetben, hogy amennyiben nem lineáris a halandóság alakulása, akkor nagyon rossz eredményeket ad. Továbbá fontos hangsúlyozni, hogy a neurális hálózat alapú hosszú távú halandósági előrejelzések területe még közel sincs kikutatva, emiatt is fontos az előrejelzések hosszú távú vizsgálata, hogy olyan módszerek kialakítására nyíljon lehetőség, amik ezeket az eredményeket stabillá és pontosabbá teszik.

Hivatkozások

- Lee, R., & Carter, L. (1992). Modeling and Forecasting U. S. Mortality. *Journal of the American Statistical Association*, 659-671.
- Altrichter, M., Horváth, G., Pataki, B., Strausz, G., Takács, G., & Valyon, J. (2006). *Neurális hálózatok*. Budapest: Hungarian Edition Panem Könyvkiadó Kft.
- Bálint, L. (2016). Mennyire illeszkedik a magyar halandóság alakulása az epidemiológiai átmenet elméleteihez? *Demográfia*, 59(1), 5-57.
- Bálint, L., & Kovács, K. (2018). Halandóság. *Demográfiai Portré*, 8, 147-175.
- Banyár, J. (2003). *Életbitosítás*. Budapest: Aula Kiadó.
- Brouhns, N., Denuit, M., & Vermunt, J. (2002.). A Poisson Log-Bilinear Regression Approach to the Construction of Projected Life Tables. *Insurance: Mathematics and Economics*, 31.(3.), 373-393.
- Caselli, G., Meslé, F., & Vallin, J. (2002). Epidemiologic transition theory exceptions. *Genus*, 58(1), 9-51.
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (old.: 1724–1734.). Doha: Association for Computational Linguistics.
- Chollet, F. (2015). *Keras*. Forrás: <https://github.com/fchollet/keras>
- Cockerham, W. (1999). Health and Social Change in Russia and Eastern Europe. *Routledge*.
- Havlik, M. (2016). Mesterséges neurális hálók modelljének javítása. *TDK DOLGOZAT*. Budapest: Budapesti Műszaki és Gazdaságtudományi Egyetem.
- HMD. (2024. március 25.). *Human Mortality Database*. Letöltés dátuma: 2024. március 25., forrás: Human Mortality Database: Hungary: <https://www.mortality.org/>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Comput*, 9.(8.), 1735–1780.

- Hyndman, R. (2023). *demography: Forecasting Mortality, Fertility, Migration and Population Data*. Forrás: R package version 2.0, <https://CRAN.R-project.org/package=demography>
- Józan, P. (2003a). Fordulat a halandóságban és az életkilátásokban, epidemiológiai korszakváltás Magyarországon? *Orvosi Hetilap*, 144(10), 451-460.
- Kingma, D., & Ba, J. (2015). Adam: A method for stochastic optimization. *ICLR*.
- Magyar Közlöny. (2022). A Magyar Nemzeti Bank elnökének 55/2022. (XII. 13.) MNB rendelete egyes biztosítási tárgyú MNB rendeletek módosításáról., 205., old.: 8798-8799.
- Májér, I., & Kovács, E. (2011). Élettartam-kockázat — a nyugdíjrendszerre nehezedő egyik teher. *Statisztikai Szemle*, 89.(7-8.), 790-812.
- Nigri, A., Levantesi, S., Marino, M., Scognamiglio, S., & Perla, F. (2019). A Deep Learning Integrated Lee-Carter Model. *Risks*, 7(33). Forrás: <https://doi.org/10.3390/risks7010033>
- Olah, C. (2015. augusztus 27). *Understanding LSTM Networks*. Letöltés dátuma: 2024. március 24., forrás: Colah's Blog: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Omran, A. (1971). A theory of the epidemiology of Population change. *Milbank Quarterly*, 83(1), 731-757.
- Omran, A. (1977). A century of epidemiologic transition in the United States. *Preventive Medicine*. *Preventive Medicine*, 6(30), 30-51.
- Omran, A. (1983). The epidemiologic transition theory. A Preliminary update. *Journal of Tropical Pediatrics*, 29(6), 305-316.
- Omran, A. (1998). The epidemiologic transition theory revisited thirty years later. *World Health Statistics Quarterly*, 51(2-4), 99-119.
- Richman, R., & Wüthrich, M. V. (2019). Lee and Carter go Machine Learning: Recurrent Neural Networks. *Swiss Association of Actuaries SAV*.
- Roell, J. (2017. június 12). *From Fiction to Reality: A Beginner's Guide to Artificial Neural Networks*. Letöltés dátuma: 2024. március 2., forrás: Medium: Towards Data Science: <https://towardsdatascience.com/from-fiction-to-reality-a-beginners-guide-to-artificial-neural-networks-d0411777571b>

- Szemenyei, M. (2021. február 8). Számítógépes Látórendszerek. *Jegyzet*.
- Szentkereszti, G., & Vékás, P. (2022). Magyar halandósági ráták előrejelzése visszacsatolt neurális hálózatokkal. *Statisztikai Szemle*, 100(10), 905-922.
- Vékás, P. (2017). Nyugdíjcélú életjáradékok élettartam-kockázata az általánosított korcsoport-időszak-kohorsz modellkeretben. *Statisztikai Szemle*, 95.(2.), 139-165.
- Vékás, P. (2019). *Az élettartam-kockázat modellezése*. Budapest: Budapesti Corvinus Egyetem.
- Vékás, P. (2022. május 11). *Bevezetés a mesterséges neurális hálózatokba*. BCE Operációkutatás és Aktuáriustudományok Tanszék, Budapest, Pest, Magyarország.
- Vékás, P. (2023. október 18). *Halandósági táblák statisztikája*. BCE Operációkutatás és Aktuáriustudományok Tanszék, Budapest, Pest, Magyarország.