

**IDŐSOR SZINKRON VALIDITÁS ELEMZÉS  
KULCSSZÓKUTATÁSON ALAPULÓ NEURÁLIS HÁLÓ  
ÁLTAL ELŐREJELZETT OLAJÁRFOLYAM  
VÁLTOZÁSON**

**TIME SERIES SYNCHRONOUS VALIDITY ANALYSIS OF  
OIL EXCHANGE RATE CHANGES PREDICTED BY  
NEURAL NETWORK BASED ON KEYWORD RESEARCH**

**BARTA Ákos – MOLNÁR Márk – NAÁRNÉ TÓTH Zsuzsanna**

**Kulcsszavak:** *neurális háló, idősor szinkron, olaj, kulcsszó kutatás, előrejelzés*

**Keywords:** *neural network, synchronous validity, oil, keyword research, forecast*

**JEL Kód:** C53

<https://doi.org/10.33565/MKSV.2023.03.01>

## **ABSZTRAKT**

*Az olaj az egyik legfontosabb nyersanyag a világgazdaságban. Szinte minden szegmensben használják, logisztika, termelés, szolgáltatások eljuttatása vagy szolgáltató eljutása a szolgáltatási helyre stb. Egyszóval a teljes gazdaságot behálózza, ezért nagyon érzékenyen reagál a piac minden volatilitásra, áremelkedésre. A nagy felhasználók esetében egyértelmű a fontosság, de a lakossági szegmensben is nagy befolyással bír, mivel a fogyasztói kosár elemeit érinti, valamint a munkabajárási költséget változtatja stb. Így az aktorok számára fontos, hogy készülni tudjanak a változásra, amennyiben nagyobb kilengés van, úgy a gazdasági döntéseiket ennek fényében tudják meghozni, így elkerülve a kiszolgáltatottságot, valamint minimalizálni a volatilitásból eredő veszteségeket. Az olajforgalmazók, pontosabban az előállítók oligopol piacon működnek, ahol több esetben egyeztetik döntéseiket, változtatnak kitermelési mennyiségein, ami fokozza az előrejelzés bizonytalanságát. Vizsgálatunk folyamán vezető gazdasági folyóiratok cikkeit elemezzük kulcsszó kutatással, illetve több egyéb változót alkalmazunk a Neurális Hálóval történő előrejelzés során, majd az így kapott idősor együttmozgását hasonlítjuk össze a valós olajár volatilitásával, ezzel felbecsülve, hogy mennyire előrejelezhetőek a változások nem bizalmi információval rendelkező szereplők számára. Az idősor elemzést több módszerrel folytatjuk le, ezzel elkerülve a részrehajlást, illetve következtetve arra, hogy milyen módon alkalmas ez a rendszer a valós és nem feltétlenül várható piaci mozgások időben való előrejelzésére, így a gazdasági döntéseink időzítésére, hogy minél kisebb legyen az ebből fakadó negatív hatás.*

## **ABSTRACT**

*Oil is one of the most important raw materials in the world economy. It is used in almost every segment, logistics, production, delivery of services or delivery of service providers to the service location, etc. In short, it involves the entire economy, which is why the market reacts very sensitively to any volatility or price increase. In the case of large users, the importance is clear, but it also has a great influence in the residential segment, as it affects the elements of the consumer basket, as well as changes the cost of commuting to work, etc. Thus, it is important for actors to*

*be able to prepare for change, if there is a larger fluctuation, they can make their economic decisions in light of this, thus avoiding vulnerability and minimizing losses resulting from volatility. The oil distributors, or more precisely the producers, operate in an oligopoly market, where in many cases they coordinate their decisions and change their production volumes, which increases the uncertainty of the forecast. In the course of our study, we analyse the articles of leading economic journals by keyword research, and we apply several other variables during the forecasting with the Neural Network, and then we compare the co-movement of the resulting time series with the volatility of the real oil price, thus estimating how predictable the changes are for actors with non-confidential information. The time series analysis is carried out using several methods, thereby avoiding partiality, and inferring how this system is suitable for predicting real and not necessarily expected market movements in time, thus for timing our economic decisions, so that the resulting negative impact is as small as possible.*

## **BEVEZETÉS**

A közgazdaságtanban a normális elvárások határain kívül eső, így előre nem láthatónak ítélt eseményeket faresteménynek nevezik (Nordhaus, 2011). Különböző üzleti alkalmazásokban kulcsfontosságú annak előrejelzése, hogy egy adott változó egy rögzített várható maximális értéket túllép az időben változó valószínűséggel. A közgazdasági kutatások hagyományosan a gazdasági mutatók előrejelzésére összpontosítanak, hogy előre jelezzék a lehetséges problémákat, kihasználják a lehetőségeket, fenntartsák a gazdasági stabilitást és csökkentsék a volatilitást. A BVAR és FAVAR modelleket sokan alkalmazták gazdasági mutatók előrejelzésére, változó sikerrel, különösen a fejlődő vagy robusztus gazdaságok esetében. (Madhou et al., 2019; Langcake & Robinson, 2017) Az egyre összetettebb modellek folyamatos fejlesztése a Big Data menedzsment adta lehetőségeknek köszönhetően több változó elemzését tette lehetővé, gyorsabb és hatékonyabb eredményt hozva. (Gupta & Kabundi, 2011)

Tekintettel arra, hogy a gazdasági eseményeket gyakran befolyásolják a kormányzati politikák, politikai események, vállalati döntések, lehetőség szerint átfogó előrejelzések készítése a cél. Például a központi bankok az inflációs ráták pontos előrejelzéseire hagyatkoznak, hogy megalapozott monetáris politikai döntéseket hozzanak, amelyek megakadályozhatják a gazdaság túlmelegedését. Hasonlóképpen, az árfolyam egy bizonyos küszöb alá süllyedésének valószínűségének előrejelzése valószínűsíthető valutaerősödést jelezhet, és megfelelő döntések meghozatalára ösztönözheti a döntéshozókat (Kumar et al., 2003).

A szakirodalom átfogó elemzéséből kiderül, hogy jelentős kutatások folytak olyan gazdasági tényezők előrejelzésére, mint az olaj- és gázárak, a kamatlábak, valamint ezek iránytrendjei. A leghatékonyabb és legkiemelkedőbb előrejelzési módszerek közé tartoznak a neurális hálózatok (Moshiri, 2000; Thakur et al., 2015; Onimode et al., 2015; Mahdiani & Khamehchi, 2016), a strukturális modellek és idősorok (Moshiri, 2000), a feltételes legkisebb négyzetek (CLS) módszere (Dadgar et al., 2006), a Bayes-féle vektorautoregresszív (BVAR) módszerek (Heidari & Parvin, 2009), dinamikus mesterséges neurális hálózatok (Naderi et al., 2018), a genetikai algoritmus (Fan et al., 2008), a wavelet elv hibrid neurális hálózattal (Jammazi & Aloui, 2012), valamint az autoregresszív modellek (Onimode et al., 2015).

A meglévő szakirodalmat áttekintve nyilvánvalóvá válik, hogy az előrejelzési módszerek két fő kategóriába sorolhatók: ökonometria és mesterséges intelligencia technikák.

A hatékony piacokon a befektetők az új információk alapján folyamatosan frissítik döntéseiket, és ennek megfelelően módosítják eszközpozícióikat. Az új információk eszközárakra gyakorolt hatása azonban az információ típusától függ. Fama és French (1993, 1996, 1997) azt találták, hogy bizonyos információk minden részvényre és azok hozamára hatással vannak, míg más információk inkább cég-specifikusak, és csak bizonyos részvények vagy alapok hozamát érintik.

(Daniel & Titman, 1997) Ezenkívül a befektetői hangulatot nem gazdasági tényezők is befolyásolhatják, mint például az időjárás. Például Kamstra és mtsai. (2003) felfedezte, hogy a piaci hozamok általában alacsonyabbak ősszel és télen, amikor az időjárás borongós, ami egy sajátos viselkedési zavart tükröz, amely a csökkent nappali órákhoz kapcsolódik.

Nincs egyetértés abban, hogy milyen típusú információk befolyásolják az eszközárakat, de a többlethozamot kereső befektetők minden információra reagálnak, és eltérően reagálhatnak ugyanarra az információhalmazra. DeLong et al. (1990) egy alapvető tanulmányában azt állítja, hogy a befektetői hangulat, amelyet úgy határoznak meg, mint a vállalat jövőbeli cash flow-ira és kockázati szintjére vonatkozó hiedelmet, amelyet a jelenlegi tények nem támasztanak alá, hosszú ideig fennmaradhat, ami olyan eszközárakhoz vezethet, amelyek jelentősen eltérnek a belső értéküktől. érték. Ennek eredményeként az eszközárak magasabbak lehetnek a befektetők túlreagálása miatt, vagy alacsonyabbak a befektetők alulreakciója miatt. (Barberis et al., 1998) Így a befektetői hangulatot tekintik a megfigyelt eszközárak szabálytalanságok egyik fő mozgatórugójának. Swaminathan (1996), valamint Neal & Wheatley (1998) a zárt végű alapok kedvezményeit a befektetői hangulat példaként tárgyalja. Neal és Wheatley (1998) továbbá azt állítja, hogy a nettó befektetési alapok visszaváltásai előre jelezhetik a téves árazás mértékét. Baker és Wurgler (2007) átfogó listát ad a befektetői hangulat tesztjeiről, többek között olyan praktikákat, mint az összesített előrejelzések, a fogyasztói bizalom változásai, a kereskedési volumen változásai és a bennfentes kereskedelem.

A mesterséges neurális hálózatok (ANN-ok) használata a gazdasági folyamatok előrejelzésében az utóbbi időben népszerűvé vált. Például Tkacz (2001) Kanada GDP-jét viszonylag kis hibaszázalékkal jósolta meg több változó segítségével, Alaminos et al. (2020) neurális hálózatot használt az inflációs ráta előrejelzésére, Galeshchuk és Demazeau (2017) pedig konvolúciós neurális hálót használt a

magyar forint árfolyamának jövőbeni értékeire. Utóbbi esetben azonban a kormányzati döntések kiszámíthatatlansága jelentősen befolyásolta az árfolyam alakulását, ami kihívást jelentett az előrejelzésre. Ennek megfelelően az ANN-ok szövegek elemzésére, következtetések levonására és a gazdasági folyamatok előrejelzésére szolgáló módszertana tovább fejlődik, és pontosabb előrejelzéseket ad. (Subecz, 2019)

A tanulmány célja, hogy felmérje az online szakmai magazinokban megjelent cikkek értékét, amelyek a tőzsdével és a gazdasággal foglalkoznak, konkrét árutőzsdei termékekkel összefüggésben. Konkrétan az ezekben a szakmagazinokban megjelent hírcikkek és a kőolaj árának változásai közötti kapcsolatot vizsgáljuk, azzal a céllal, hogy előre jelezzük a jövőbeli árakat.

## **KUTATÁSMÓDSZERTAN**

A kutatás felépítése több lépésből áll. Első körben taglaljuk az olajpiacot, vizsgáljuk az aktorok szempontjából, valamint indokoljuk, hogy egyes aktorok döntései miért olyan fontosak. Majd adatgyűjtés történik, úgynevezett Big Data adathalmazt készítünk scraper segítségével. (Scrapelés: web scraping, magyar fordítás „web kaparás”, de ez nem használatos. Online, többnyire adatgyűjtés a weboldalak html szerkezetéből.) Az adatokat Neurális Hálóval tanító és adathalmazokra bontjuk, ezek segítségével alkotjuk meg a hálót. Majd az eredményeket idősorosan vizsgáljuk együttmozgás szempontjából.

Az olajpiac kevés szereplős, így egyes szereplőknek nagy hatásuk van a piaci mozgásokra. (1. táblázat)

## 1. táblázat. Az olajtermelés megoszlása világszerte 2021-ben

Közel-Kelet	Észak-Amerika	Európa és CIS	Ázsia	Afrika	Dél- és Közép Amerika
31,3%	26,6%	19,3%	8,4%	7,8%	6,6%

*Forrás: statista.com, 2018*

Nem feltétlenül tudnak érvényesülni a szabadpiaci mozgások, hiszen a szereplők erősen egymás döntéseire utaltak, vannak vezető és követő aktorok is. Egy ilyen piaci működésben az előrejelzések, döntési folyamatok ismerete nagyfokú biztonságot adhat a gazdaság szereplői számára, hogy felkészüljenek az esetleges árvolatilitásra. Vagyis egy olyan piacról beszélhetünk, ahol a termék vagy nyersanyag ára a kevés disztribútor miatt nem csak és kizárólag a kereslet-kínálat függvényében változik. Vagyis pontosabban aszerint, viszont erre az aktorok döntései nagy hatással bírnak, így manipulálhatják.

Kutatásunkban azt feltételezzük, illetve azt vizsgáljuk, hogy feltételezések szerint több tőzsdei mozgás, volatilitás és trendforduló spekulációs alapokon nyugszik. Nem feltétlenül teljes mértékben spekulatív, de jelentős szerepet vállal az árfolyam változásában a hiedelmek és valószínűségszámításokon, becsléseken alapuló jövőbeli árfolyamelőrejelzés. Mivel a döntéshozók gondolatait nem ismerhetik, így hiteles forrásokat keresnek, jelen esetben a vezető gazdasági folyóiratokat tekintjük hiteles forrásnak. Vagyis azt kutatjuk, hogy azon forrásokat, amiből spekulatív jelleggel döntéseket alapoznak, mennyire validan van összeköttetésben az azokat követő árfolyamváltozásokkal. Így Neurális Hálóval igyekszünk következtetéseket és konvergenciakapcsolatokat levonni.

A nagy adathalmaz elemzése során bizonyos statisztikai modellek is célravezetőek lehetnek, mindamelllett egy Neurális Háló több ezer kapcsolatot, összefüggést és következtetést von le. A nagy adathalmaz, a rengeteg egymásrataltság és

összekapcsolt mutatók idősoron átívelése miatt a legoptimálisabb és leghatékonyabb módszer egy Neurális Háló felépítése az adott projektre.

A Big Data felépítése során szükség volt egy olyan folyóíratra volt szükség, amely kellően olvasott, hogy véleményformáló legyen, időben tudósít fontos, olajpiacot érintő eseményeket, nyilatkozatokat.

Összességében a folyóiratnak az alábbi szempontoknak kellett megfelelni:

- vezető folyóirat, nagy olvasói számmal, vagyis tényleges ráhatása lehet a piacok működésére, spekulatív döntéshozatalra,
- rendelkezik archív állománnyal, mely hozzáférhető, kellően strukturált,
- scrapelhető, bár ez inkább programozási kérdés.

Több folyóirat (Forbes, Businessinsider, Wall Street Journal, Bloomberg, Reuters, Yahoo! finance, CNBC, NYTimes) alapos vizsgálata és elemzése során a Wall Street Journal-ra (továbbiakban: WSJ) esett a választás. Ez elég markáns folyóirat, de az elemzettek között volt hasonló vagy talán markánsabb. A döntő faktor az archívum és a scrapelhetőség volt. Több folyóirat nem rendelkezett archívummal vagy adott esetekben nagyon rendezetlen volt, így összességében jelen projektre a WSJ volt a legalkalmasabb.

A Mesterséges Neurális Hálók (Artificial Neural Network, továbbiakban: NN), egy agyi neurológia hálózathoz hasonló adatfeldolgozó rendszer. A neuronok összeköttetésben vannak egymással, az összekötések, kapcsolatok különböző súlyokkal rendelkeznek, melyeket a tanítás során optimalizálunk, ezzel létrehozva egy pontos és hatékony NN-t.

Az NN tanítása során egy tanítóhalmazt választunk le az adattömegeből, melyen végigfuttatjuk az NN-t. A súlyokon módosítunk, optimalizálunk, majd újra futtatjuk. Ezt addig folytatjuk, amíg a mért hibahatár standardizálódik, vagyis egy adott szám körül ingadozik. Ebben az esetben további futtatásokkal eredményt javítani nem tudunk. Hiba a legtöbb esetben van. Illetve, mivel nem egzakt dolgot vizsgálunk, hanem vélelmezett emocionális kapcsolatot, így elméletileg ilyen

helyzetben nem létezhet olyan neurális háló, amely hiba, másszóval eltérés nélkül képes lefutni.

A kutatás során a Wall Street Journal cikkekben specifikus kulcsszavak alapján végeztünk elemzéseket, hogy meghatározzuk, mennyire tárgyalják az olajárat és annak változásait a cikkekben. A vizsgált kulcsszavak között olyanok szerepeltek, mint 'crude', 'opec', 'oil price', 'wti', 'increase', 'rise', 'rising', 'grow', 'optimism', 'enhance', 'expensive', 'climb', 'optimal', 'agreement', 'cooperation', 'solution', 'deal', 'bull', 'gain', 'demand', 'positive', 'decrease', 'bear', 'fall', 'low', 'cut', 'dramatically', 'pessimism', 'emergency', 'emerge', 'recession', 'collapse', 'negative', 'reduce', 'disagree', 'decline', és 'cheap'.

Különös hangsúlyt fektettünk az első négy kulcsszóra: 'crude', 'opec', 'oil price', és 'wti'. Ezeket az indikátorként használtuk, melyek meghatározták, hogy a cikk valószínűleg az olajárról szól-e vagy sem. Csak azokat a cikkeket elemeztük részletesen, amelyekben ezek az indikátorok megfelelő gyakorisággal fordultak elő.

A kulcsszavak elemzésének kiegészítéseként a cikkek hangulati, avagy sentiment elemzését is elvégeztük az nltk vader lexicon segítségével. A VADER (Valence Aware Dictionary and sEntiment Reasoner) lexicon egy kifejezetten közösségi média tartalmak hangulati elemzésére kifejlesztett eszköz. A szentiment elemzés során a szövegekben található kifejezések pozitív, negatív vagy semleges érzelmeket fejeznek-e ki. A VADER lexicon előnye abban rejlik, hogy képes kezelni az internetes nyelvezet és szleng sajátosságait, így például az iróniát vagy az emotikonokat is. Az nltk (Natural Language Toolkit) egy Python nyelven íródott könyvtár, amely számos eszközt és angol nyelvű korpuszt tartalmaz a számítógépes nyelvészeti kutatásokhoz. Az nltk VADER modulja ezt a kifejezetten hangulatelemzésre szánt lexikont használja, így teszi lehetővé a felhasználók számára, hogy egyszerűen és hatékonyan végezzenek sentiment elemzést szövegeken. Ezen elemzés eredményeként képesek voltunk

meghatározni a cikkekben kifejezett pozitív vagy negatív hozzáállást az olajárral és annak változásaival kapcsolatban. Az összes kulcsszó és sentiment elemzés kombinációjával mélyreható képet kaptunk arról, hogy milyen kontextusban és milyen érzelmekkel tárgyalják az olajárakat a Wall Street Journalban.

A kapott eredményeket sorba rendeztük, majd összevetettük az olajár idősoros eredményeivel az alábbi elemzések felhasználásával:

- Pearson korreláció
- Time Lagged Cross Correlation
- Dynamic Time Warping

A Pearson korreláció, a Time Lagged Cross Correlation (TLCC) és a Dynamic Time Warping (DTW) három különböző, de komplementer elemzési módszer, melyek különösen alkalmasak lehetnek a Wall Street Journal cikkeiben szereplő kulcsszavak és az olajár napi változásának összekapcsolására, ha előrejelzéseket szeretnénk készíteni. Az alábbiakban ismertetjük, hogy miért:

Pearson korreláció:

- Mértékadó: A Pearson korreláció az egyik legelterjedtebb statisztikai módszer, amely képes mérni két változó közötti lineáris kapcsolat erősségét és irányát. A kulcsszók előfordulásának és az olajár változásának közötti potenciális összefüggések azonnali felismerésére használható.
- Egyszerűség: Egyszerű értelmezni és implementálni, ezért ideális kiindulópont lehet az elemzések során.
- Limitációk kezelése: Bár csak lineáris összefüggéseket képes észlelni, ez egy gyors módja annak, hogy felfedezzük, van-e egyáltalán kapcsolat a két változó között.

Time Lagged Cross Correlation (TLCC):

- Időbeli eltérések figyelembevétele: A TLCC lehetővé teszi számunkra, hogy meghatározzuk, mennyi idővel később hatnak a kulcsszavak az olajár

változására. Ez különösen fontos, mert az információ terjedésének és a piac reakciójának időbeli eltolódása lehet.

- Dinamikus változások: Meg tudja mutatni, hogy a kapcsolat erőssége az idő során hogyan változik, amely kritikus lehet a változó piaci körülmények között.

Dynamic Time Warping (DTW):

- Nem lineáris összehasonlítás: A DTW lehetővé teszi két idősor összehasonlítását, még akkor is, ha azok időbeli eltolódásokat vagy szakaszos nyújtásokat mutatnak. Ez különösen hasznos lehet, ha az olajár változásai és a kulcsszavak nem ugyanazon a sebességen, vagy nem ugyanabban az időpontban következnek be.
- Rugalmas: A DTW több információt tud nyújtani a változások dinamikájáról, és jobb lehet a nem lineáris kapcsolatok azonosításában, mint a Pearson korreláció.

Amikor előrejelzéseket készítünk a Wall Street Journal cikkeiben szereplő kulcsszavak és az olajár napi változásának összefüggésére, fontos, hogy különböző nézőpontokból közelítsünk a problémához, és figyelembe vegyünk a különböző időbeli és nem lineáris összefüggéseket. A Pearson korreláció, a TLCC és a DTW kombinálása ezt a sokoldalú megközelítést biztosítja, ami növelheti az előrejelzések pontosságát és megbízhatóságát.

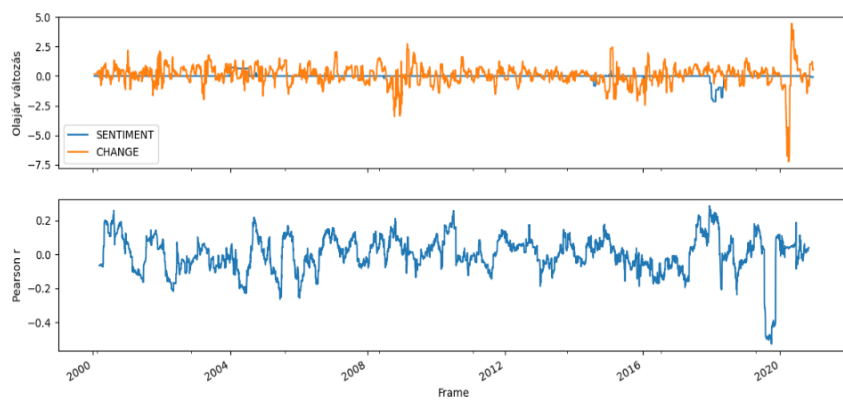
## **EREDMÉNYEK**

A Pearson-korreláció lineáris kapcsolatot mér két folytonos jel között, -1 és 1 közötti értékekkel, ahol -1 negatív, 1 pozitív korrelációt, 0 pedig semmilyen korrelációt jelöl. Hátránya, hogy csak lineáris összefüggést képes detektálni és kiugró értékek torzítják az eredményeket.

Különböző idősorok szinkronizálásának elemzése során különféle módszerekkel találkozhatunk, melyek eltérő előnyökkel és hátrányokkal rendelkeznek. A

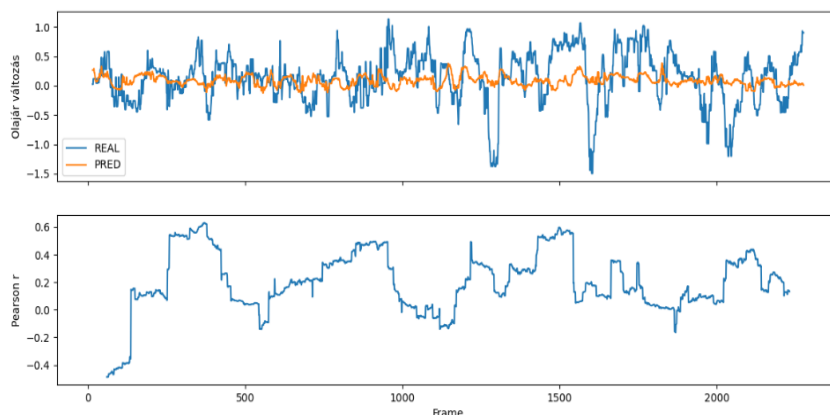
Pearson korreláció egyszerű és értelmezhető, de csak lineáris kapcsolatokat mér, míg más módszerek, mint az időkélesztetett keresztkorreláció vagy a dinamikus idővetemítés, időbeli késleltetéseket és nem lineáris összefüggéseket detektálnak, de nehezebb értelmezni és számításigényesek lehetnek.

Egy példán keresztül megvizsgálva, az WSJ cikkek hangulatelemzésekor alkalmazott Pearson-korreláció adatai közt gyakorlatilag nincs lineáris összefüggés, amit a 0.0091-es  $r$  érték és a 0.42  $p$ -érték is alátámaszt. (1. ábra)



**1. ábra. Olajár változás és WSJ Hangulatelemzés Pearson korreláció**  
*Saját szerkesztés*

Az ANN előrejelzésű modellnél, ahol csak a tesztalmez eredményeit vettük figyelembe, gyenge pozitív lineáris összefüggést találtam a változók között, amit a 0,115-ös  $r$  érték és a 0.003008-as  $p$ -érték igazol. Ez azt jelenti, hogy valószínűleg van összefüggés a változók között, de további vizsgálatokra van szükség az okok meghatározásához. (2. ábra)



**2. ábra. Olajár változás és ANN előrejelzés Pearson korreláció**  
*Saját szerkesztés*

Az időképletetett keresztkorreláció (TLCC) módszere lehetőséget ad arra, hogy az idősorok közötti kapcsolatokat az időben történő késleltetések figyelembevételével vizsgáljuk. Ez különösen hasznos például kauzális kapcsolatok megértésében, hiszen megállapíthatjuk, hogy az egyik idősor hogyan befolyásolja a másikat bizonyos időbeli késleltetéssel. A módszer hátránya viszont, hogy nagy számú késleltetés esetén a számítások sokkal bonyolultabbakká válnak, és a korrelációk értelmezése nehezebb lehet.

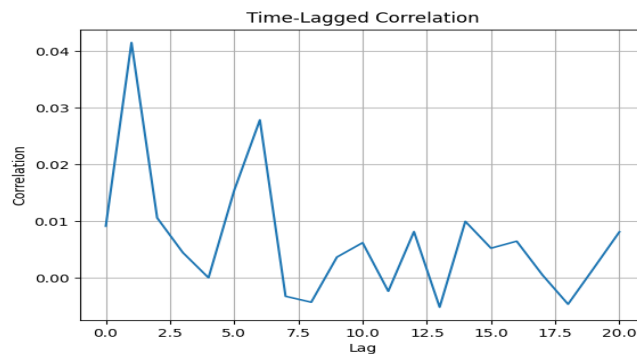
A TLCC módszert a Wall Street Journal cikk hangulatelemzésére alkalmaztuk. Ennek során a "max lag" kifejezést használtuk, mely kifejezi, hogy a két idősoros adatsor közötti korrelációt kiszámításához mekkora a maximális eltolás az időben. Az eltoláshoz kapcsolódó korrelációt a legkisebb (0, azaz nincs eltolás) és a legnagyobb (az előre meghatározott max lag) érték között találjuk, hogy azonosítsuk, melyik eltolási érték mellett a legerősebb a kapcsolat a két idősor között.

A vizsgált változók közötti korreláció értékeinek minőségi értelmezése is fontos. Ha a korreláció értéke 1 (praktikusan tökéletes pozitív lineáris kapcsolat), az azt mutatja, hogy az idősorok együttesen növekednek vagy csökkennek. Ha viszont

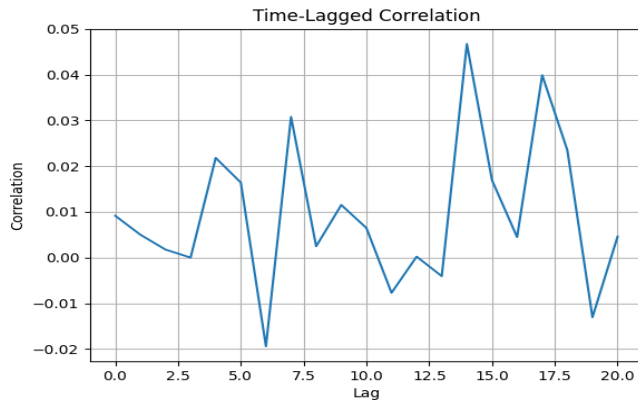
az érték alacsony, mint például 0,04, az csak gyenge pozitív lineáris kapcsolatot jelez, vagyis a két idősor nem igazán mozog együtt, és a közöttük lévő kapcsolat gyenge.

A TLCC elemzés során az  $x$  értékét annak az időornak állítjuk be, melyiket előbb szeretnénk elemezni. Ezen azonban nincs szigorú szabály, de általánosságban az az idősor kerül beállításra  $x$ -nek, amelynek eseményei előbb következnek be. Ennek ellenére, a korreláció szimmetrikus, tehát a kapcsolat előjele változik, ha az  $x$  és  $y$  idősorok helyét felcseréljük, de az abszolút értéke nem.

A vizsgálatokat kiterjesztettük az ANN által előrejelzett eredményekre is. (3., 4. és 5. ábra) Az eredmények azt mutatták, hogy a legnagyobb lineáris kapcsolat akkor mutatkozik, amikor nincs időbeli eltolás a két idősor között, azonban az érték még mindig csak gyenge pozitív összefüggésre utal. Ez arra utal, hogy a kapcsolat a két idősor között nem erős, de mégis van bizonyos mértékű összefüggés azok között.

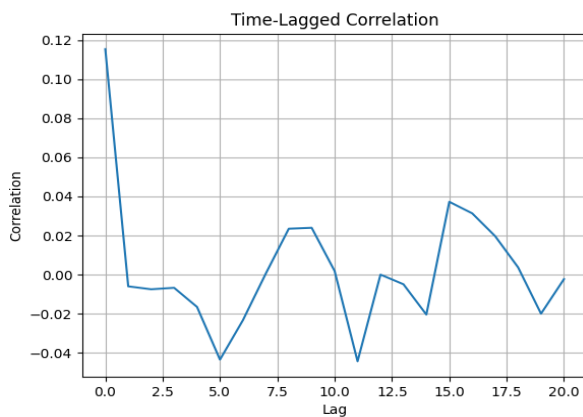


**3. ábra. TLCC eredmények olajár változás és hangulatelemzés**  
**( $x$ =Hangulat)**  
*Saját szerkesztés*



4. ábra. TLCC eredmények olajár változás és hangulatelemzés (x=Olajár)

*Saját szerkesztés*

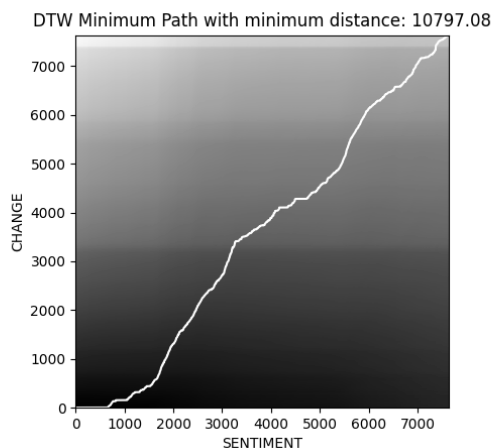


5. ábra. TLCC eredmények olajár változás és ANN eredmények (x=ANN)

*Saját szerkesztés*

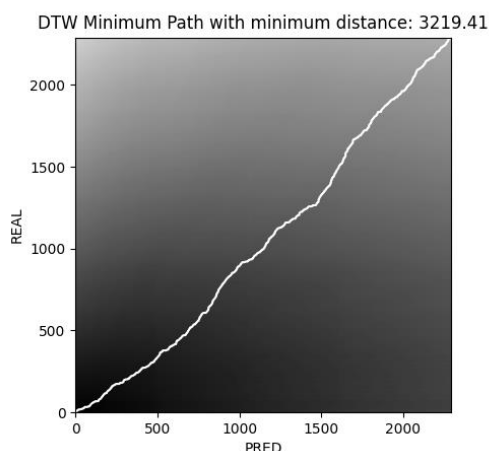
A Dynamic Time Warping (DTW) egy hatékony eszköz idősorok összehasonlítására, amely figyelembe veszi az időbeni nyújtásokat és összehúzóásokat, így nemlineáris összefüggések detektálására is képes. Előnye a rugalmassága és az alkalmazhatósága különböző idősorok esetén, de nagy adathalmazok esetén a számítási igényessége miatt nehezebb használni.

DTW eredmények értelmezésének szabályai szerint a két idősor közötti hasonlóságot jelző érték: ha 0-hoz közeli, a két idősor nagyon hasonló; ha az érték növekszik, a két idősor közötti eltérés is nő. Így a kisebb DTW értékkel rendelkező idősor-párok hasonlóbbak egymáshoz.



**6. ábra DTW eredmények olajár változás és hangulatelemzés**  
*Saját szerkesztés*

Olajár változás és WSJ hangulatelemzés közötti DTW elemzés eredménye 10797.08 volt, ami jelentős eltérésre utal, tehát a két idősor kevésbé hasonlít egymáshoz. (6. ábra) Ez arra utalhat, hogy a hangulatelemzés és az olajár változása nem követik szorosan egymást. Ezzel szemben, az olajár változás és az ANN által előrejelzett olajár közötti DTW érték csak 3219.41, ami kisebb eltérést és nagyobb hasonlóságot mutat. (7. ábra)



**7. ábra. DTW eredmények olajár változás és ANN eredmények**  
*Saját szerkesztés*

## KÖVETKEZTETÉSEK

A kutatásunk eredményei összhangban állnak a neurális hálózatokkal kapcsolatos korábbi kutatásokkal, melyek kimutatták az NN-k hatékonyságát komplex összefüggések modellezésében. Különösen a pénzügyi piacokra és az olajárfolyam előrejelzésére vonatkozó kutatásokban számos példa található erre.

A korrelációs elemzések kapcsán kijelenthető, hogy összefüggés létezik az NN által előrejelzett eredmény időSORa és a valós olajárváltozás időSOR között. A kapcsolat helyenként erős, helyenként viszont elhanyagolható.

Pearson: az összefüggés egyértelműen kimutatható, a kapcsolat nem túl erős

TLCC: a változók között a kapcsolat relatíve gyenge

D'TW: a kapott D'TW eredmények alapján a hangulatelemzés és az olajár változása között kevésbé valószínű az erős összefüggés, az olajár változása és az ANN előrejelzése között viszonylag nagy a hasonlóság.

Fontos megjegyezni, hogy míg a Pearson korreláció és a TLCC gyengébb eredményeket mutatott a jelen kutatásban, korábbi vizsgálatokban gyakran ezek a módszerek nyújtottak nagyobb mértékű korrelációt. Ezt a különbséget az adatok specifikus jellege vagy a vizsgált időSORok sajátosságai okozhatták.

Ami a DTW-t illeti, a jelen kutatás eredményei az olajár változásainak és az ANN előrejelzésének közötti hasonlóságokat mutatnak be. Korábbi kutatások is használták a DTW-t idősorok összehasonlításában, és megerősítették annak alkalmasságát nem lineáris összefüggések detektálására.

Kutatási szempontból megfogalmazhatjuk, hogy az előrejelzés bizonyos szintig validnak tekinthető, mindazonáltal a további hatékonyság növelés kérdéses. Azaz az NN-nel történő olajárfolyam változás előrejelzés az esetek egy jelentős részében valós, jövőre nézve jó eséllyel bekövetkező árfolyamváltozást jelzett.

A Neurális Háló által feltárt következtetések, illetve az optimalizálási folyamatban kapott eredmények alapján a módszer hatékonynak tekinthető. Mármint a módszert illetően vélhetően a maximális korrelációt vagy legalábbis a maximumhoz legközelebb eső ráhatást sikerült kimutatni.

Érdekes összevetni, hogy míg a jelen kutatás kimutatta az NN általi előrejelzések és az olajárfolyam változásai közötti jelentős kapcsolatot, néhány korábbi kutatás inkább óvatosságra intett az ilyen előrejelzésekkel kapcsolatban, különösen a pénzügyi piacok volatilitása és az információáramlás gyorsasága miatt.

Végül, a kutatásunk az NN hatékonyságát hangsúlyozza. Ezen eredmények fényében érdemes további kutatásokat végezni annak érdekében, hogy jobban megértsük, milyen körülmények között és milyen mértékben tudjuk növelni az előrejelzések pontosságát és megbízhatóságát.

Jelenlegi eredmények tekintetében azonban az eredmények nem minden kétséget kizáróak, így szükséges mérlegelni a lehetőségeket az eredmények és validitás pontosítására. (Mivel a változók között az összefüggés kimutatható, így a modell javításán vagy pontosításán szükséges és érdemes is dolgozni!) Egyrészt növelhető azzal, ha adott paraméterek függvényében nem a teljes idősort, hanem csak adott részeit vizsgáljuk az NN vagy trend függvényében, ezzel lehetséges, hogy sokkal jobban együttmozgó idősorokat kapunk. Jelen kérdés és vizsgálat lefolytatásához

szükséges, hogy meghatározzuk és teszteljük különböző perifériaváltozókra, majd az így kapott eredményeke elemezzük.

## IRODALOMJEGYZÉK

1. Alaminos, D., Esteban, I., Salas, M.B. & Callejón, A.M., 2020. Quantum Neural Networks for Forecasting Inflation *Dynamics, Journal of Scientific & Industrial Research*. (79) 103-106. <https://doi.org/10.56042/jsir.v79i2.68439>
2. Baker, M.P. & Wurgler, J.A., 2007. Investor Sentiment in the Stock Market. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.962706>
3. Barberis, N., Shleifer, A. & Vishny, R., 1998. A model of investor sentiment. *Journal of Financial Economy*. (4) 307-343.
4. Dadgar Y., Keshavarz Gh. & Tyataraj A., 2006. The Analysis of Relationship Between Inflation and Economic Growth in Iran, *Journal of Economic Literature* (5) 59-88.
5. Daniel, K. & Titman, S. (1996). Evidence on the Characteristics of Cross Sectional Variation in Stock Returns. *The Journal of Finance*, 52(1), 1-33. <https://doi.org/10.2307/2329554>
6. DeLong, J.B., Shleifer, A., Summers, L.H. & Waldmann, R.J., 1990. Noise trader risk in financial markets. *Journal of Political Economy*. (98), 703-738.
7. Fama, E.F. & French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56. [https://doi.org/10.1016/0304-405x\(93\)90023-5](https://doi.org/10.1016/0304-405x(93)90023-5)
8. Fama, E.F. & French, K.R., 1996. Multifactor Explanations of Asset Pricing Anomalies. *The Journal of Finance*. 51(1), 55-84. <https://doi.org/10.1111/j.1540-6261.1996.tb05202.x>
9. Fama, E.F., French, K.R., 1997. Industry cost of equity. *Journal of Financial Economy*. (43), 153–193.
10. Fan, Y., Liang, Q. & Wei, Y.M., 2008. A generalized pattern matching approach for multi-step prediction of crude oil price. *Energy Economics*, 30(3), 889-904. <https://doi.org/10.1016/j.eneco.2006.10.012>
11. Galeshchuk, S. & Demazeau, Y., 2017. Forecasting hungarian forint exchange rate with convolutional neural networks. *International Conference on Behavioral, Economic, Socio-Cultural Computing (BESCC)*. <https://doi.org/10.1109/besc.2017.8256358>
12. Gupta, R., & Kabundi, A. (2011). A large factor model for forecasting macroeconomic variables in South Africa. *International Journal of Forecasting*, 27(4), 1076-1088. <https://doi.org/10.1016/j.ijforecast.2010.10.001>
13. Heidari H. & Parvin S., 2009. Modeling and Forecasting Iranian Inflation with Time Varying BVAR Models, *Iranian Journal of Economic Research* (36) 59-84.

14. Jammazi, R. & Aloui, C., 2012. Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling. *Energy Economics*, 34(3), 828-841. <https://doi.org/10.1016/j.eneco.2011.07.018>
15. Kamstra, M.J., Kramer, L.A. & Levi, M.D., 2000. Winter Blues: A SAD Stock Market Cycle. *SSRN Electronic Journal* <https://doi.org/10.2139/ssrn.208622>
16. Kumar, M., Moorthy, U. & Perraudin, W., 2003. Predicting emerging market currency crashes. *Journal of Empirical Finance*, 10(4), 427-454. [https://doi.org/10.1016/s0927-5398\(02\)00068-3](https://doi.org/10.1016/s0927-5398(02)00068-3)
17. Langcake, S. & Robinson, T., 2017. Forecasting the Australian economy with DSGE and BVAR models. *Applied Economics*, 50(3), 251-267. <https://doi.org/10.1080/00036846.2017.1319558>
18. Madhou, A., Sewak, T., Moosa, I. & Ramiah, V., 2019. Forecasting the GDP of a small open developing economy: an application of FAVAR models. *Applied Economics*, 1–12. <https://doi.org/10.1080/00036846.2019.1679346>
19. Mahdiani, M.R. & Khamehchi, E., 2016. A modified neural network model for predicting the crude oil price. *Intellectual Economics*, 10(2), 71–77. <https://doi.org/10.1016/j.intele.2017.02.001>
20. Moshiri, S. & Cameron, N., 2000a. Neural network versus econometric models in forecasting inflation. *Journal of Forecasting*, 19(3), 201-217. [https://doi.org/10.1002/\(sici\)1099-131x\(200004\)19:3<201::aid-for753>3.0.co;2-4](https://doi.org/10.1002/(sici)1099-131x(200004)19:3<201::aid-for753>3.0.co;2-4)
21. Naderi, M., Khamehchi, E. & Karimi, B., 2018. Novel statistical forecasting models for crude oil price, gas price, and interest rate based on meta-heuristic bat algorithm, *Journal of Petroleum Science and Engineering*, <https://doi.org/10.1016/j.petrol.2018.09.031>
22. Neal, R. & Wheatley, S.M., 1998. Do Measures of Investor Sentiment Predict Returns? *The Journal of Financial and Quantitative Analysis*, 33(4), 523. <https://doi.org/10.2307/2331130>
23. Nordhaus, W.D., 2011. The Economics of Tail Events with an Application to Climate Change. *Review of Environmental Economics and Policy*, 5(2), 240-257. <https://doi.org/10.1093/reep/rer004>
24. Onimode, B., Alhassan, J., and Adepoju, S. (2015). Comparative study of inflation rates forecasting using feed-forward artificial neural networks and auto-regressive (ar) models. *International Journal of Computer Science Issues*, 12.
25. statista.com, Distribution of oil production worldwide in 2021, by region [online] <https://www.statista.com/statistics/277621/distribution-of-global-oil-production-by-region/> (2023.05.10.)
26. Subecz, Z., 2019. Event Detection and Classification in Natural Texts, *Gradus* 6(1), 16-21.

27. Swaminathan, B., 1996. Time-Varying Expected Small Firm Returns and Closed-End Fund Discounts. *Review of Financial Studies*, 9(3), 845-887. <https://doi.org/10.1093/rfs/9.3.845>
28. Thakur, G.S.M., Bhattacharyya, R. & Mondal, S.S., 2016. Artificial Neural Network Based Model for Forecasting of Inflation in India. *Fuzzy Information and Engineering*, 8(1), 87-100. <https://doi.org/10.1016/j.fiae.2016.03.005>
29. Tkacz, G., 2001. Neural network forecasting of Canadian GDP growth. *International Journal of Forecasting*, 17(1), 57–69. [https://doi.org/10.1016/s0169-2070\(00\)00063-7](https://doi.org/10.1016/s0169-2070(00)00063-7)