

A gépi tanulás négy lovasa a pénzügyekben

Matthew Dixon
Alkalmazott Matematikai Tanszék,
Stuart School of Business
(udvariasságból),
Illinois Institute of Technology
matthew.dixon@iit.edu

Igor Halperin
Fidelity Investments

Szeptember
15,2019

Absztrakt

A gépi tanulást már több mint 40 éve használják a pénzügyi szolgáltatási ágazatban, de csak az utóbbi években terjedt el egyre inkább a befektetési menedzsment és a kereskedés területén. A gépi tanulás általánosabb keretet biztosít a pénzügyi modellezéshez, mint lineáris parametrikus elődje, általánosítva az archetipikus modellezési megközelítéseket, mint például a faktormodellezés (Feng et al., 2018; Gu et al., 2018; Chen et al., 2019), a származtatott termékek árazása, a portfólióépítés, az optimális fedezeti ügyletek modellmentes, adatvezérelt megközelítésekkel (Halperin, 2017), amelyek robusztusabbak a modellkockázatra és a kiugró értékek rögzítésére. A bizonyított potenciáljuk ellenére azonban az elfogadás útjában álló akadályok is megjelentek - ezek többsége a tudományközi terület szociológiájának műtárgya. Ez az állásfoglalási cikk több iparági szakértővel folytatott megbeszélések és a szerzők több évtizedes, a gépi tanulást és a hagyományos kvantitatív pénzügyeket befektetési bankoknál, vagyongazdálkodó és értékpapír-kereskedő cégeknél alkalmazó tapasztalatai alapján azonosítja a főbb veszélyeztetőket, és iránymutatásokat és megoldásokat fogalmaz meg ezek elkerülésére. A legjobb gyakorlatok illusztrálására a felügyelt tanulást és a megerősítést a befektetési menedzsmentben és a kereskedésben alkalmazó példák kerülnek bemutatásra.

1 Bevezetés

A gépi tanulás gyakorlata a pénzügyekben a gépi tanulás teoretikai és számítástechnikai fejlődésével arányosan nőtt. A korai alkalmazók a kvantitatív fedezeti alapok voltak, köztük a Bridgewater Associates, a Renaissance Technologies, a Worldquant, a D.E. Shaw és a Two Sigma, akik felkarolták az új gépi tanulási technikákat, bár az elfogadás mértéke vegyes, és egészséges szkepticizmus van abban a tekintetben, hogy a gépi tanulás csodaszor a kvantitatív kereskedés számára. 2015-ben a Bridgewater Associates bejelentette, hogy új mesterséges intelligencia egységet hozott létre, miután az IBM Watson-tól mélytanulásban jártas embereket vett fel. Anthony Ledford, az MAN AHL vezető kutatója: "Ez még korai stádiumban van. Félretettünk egy potnyi pénzt a tesztkereskedésre. A mélytanulással, ha minden jól megy, a tesztkereskedésbe fog kerülni, ahogy más gépi tanulós megközelítések is". David Harding, a Winton Capital Management vezérigazgatója: "Az emberek azt kezdték mondani, hogy 'Van egy

Az ebben a tanulmányban kifejtett vélemények a szerzők sajátjai, és nem feltétlenül tükrözik munkáltatójuk véleményét.

elképesztő új számítástechnika, amely mindent elsőprő lesz, ami eddig volt. A genetikai algoritmusok is divatba jöttek. Nos, elmondhatom, hogy ma már egyik ilyen cég sem létezik.

- egy kolbász sem".

Az ML mint kutatási terület a pénzügyekben még mindig csak kialakulóban van, és fontos szempontból hiányos. Mindazonáltal a technológia legkorábbi és legaktívabb felhasználói és fejlesztői az üzleti döntéshozatalban a pénzügyek területén működtek, lásd például (Martin, 1977; Trippi és Desieno, 1992; Swanson és White, 1995). A pénzügyekben minden szükséges összetevő megvan ahhoz, hogy az ML működjön: hatalmas mennyiségű adat, számítási és emberi erőforrások, közvetlen nyereség- és veszteségkövetkezmények, valamint egy rendkívül versenyképes környezet, ahol minden előnyre szükség van a siker érdekében. Ugyanakkor a szabályozó hatóságok egyre inkább ellenőrzik, és néhány alapvető kutatási tevékenység ellenáll az "empirikus" vagy "mérnöki" megközelítésekre jellemző "fekete doboz" módszertanoknak.

Mindazonáltal a pénzügyi rendszeren belüli tevékenységek rögzítésére és kommunikálására szolgáló, gépileg olvasható adatok növekedése, valamint a számítási teljesítmény és a tárolási kapacitás folyamatos növekedése jelentős hatással van a pénzügyi modellezés minden szegletére. A 2007- 2008-as pénzügyi válságok óta a szabályozó felügyelet az "adatvezérelt" szabályozás felé orientálódott, aminek egyik kiemelkedő példája a részletes szerződési feltételek összegyűjtése és elemzése az Egyesült Államokban és Európában a válság hatására indított banki hitel- és kereskedési könyv stressztesztelési programokhoz (Flood et al., 2016).

Az "alternatív adatok" - amelyek az értékpapírok árazásának, a vállalati fundamentumoknak vagy a makrogazdasági mutatóknak a szokásos körén kívüli adatokra és információkra utalnak - egyre fontosabb szerepet játszanak az eszközkezelők, kereskedők és döntéshozók számára. A közösségi média ma már a fedezeti alapok által jelenleg használt alternatív adatok egyik legfontosabb kategóriájaként szerepel. A kereskedőcégek olyan gépi tanulással foglalkozó szakértőket alkalmaznak, akik képesek a természetes nyelvi feldolgozás (NLP) alkalmazására a pénzügyi hírekre és más strukturálatlan dokumentumokra, például a nyereségbejelentő jelentésekre és a SEC 10K jelentésekre. Az olyan adatszolgáltatók, mint a Bloomberg, a Thomson Reuters és a RavenPack a szisztematikus kereskedési modellekhez szabott, feldolgozott hír-érzelmi adatokat szolgáltatnak. Lásd Guida (2019) és de Prado (2018) a gépi tanulás és a nagy adatok kiterjedt tárgyalását a kvantitatív befektetésben és kereskedésben.

A gépek képesek modellezni az összetett és nagy dimenziójú adatgenerálási folyamatokat, több millió modellkonfiguráción átfutni, majd robusztusan értékelni és korrigálni a modelleket az új információk hatására (Dhar, 2013). Azáltal, hogy folyamatosan frissítik és számos egymással versengő modellt fogadnak, megakadályozzák, hogy bármelyik modell egy olyan adatgyűjtő silóba vezessen minket, amely csak az adott piaci nézet szempontjából hatékony. Szerkezetileg az adattudomány egyik mellékterméke az, hogy megváltoztattuk a gondolkodásmódunkat - az ML használatával az adatokból való érvelés, kísérletezés és perspektíváink alakítása robusztusabb gyakorlatokat ápol az empirikusan vezérelt kereskedési és befektetési döntési folyamatokhoz. A területnek a végső gondolkodó gép iránti megszállottsága természetéből adódóan ironikus, hogy a gépi tanulás pénzügyi alkalmazására vonatkozó negatív eredményekhez hozzájáruló tényezőknek emberi tényezőknek kell lenniük. Egy olyan területen, amely erősen mérnököközpontú, és az AI-guruk ikonizálódnak, talán nem meglepő, hogy a sikeres pénzügyi alkalmazások előállításának kontextuális kihívásai a gépi tanulásban nem mindig valósultak meg. A pénzügy annyiban egyedi terület, hogy az innováció nagy részét más tudományágaktól kölcsönzi, amelyek közül egyesek tapsolnak a nyílt innovációnak, mégis gyakran más és más, titkos módon működik. Míg a tudományos szakirodalom a gépi tanulás pénzügyi alkalmazásaival foglalkozik, és pontszerű megoldásokat kínál az egyes kihívásokra, kevés olyan anyag van, amely megközelíti a

az örökbefogadás témája holisztikus szemszögből.

Először röviden áttekintjük a gépi tanulást a pénzügyekben, majd azonosítjuk a főbb szereplőket és érvelési típusokat, amelyek komoly hibákhoz vezettek a gépi tanulási alkalmazásokban.

Ezek közül néhány nem egyedülálló a gépi tanulásban a pénzügyekben, sőt, a történelem ismétli önmagát a pénzügyek korábbi példáival, amelyek megelőzték a gépi tanulást. Ezután egy olyan elvi megközelítést mutatunk be a kísérlettervezéshez, amely ezeket a problémákat igyekszik ellensúlyozni.

2 Mi az a gépi tanulás a pénzügyekben?

A gépi tanulás a pénzügyekben az adatokból történő pénzügyi modellezésre szolgáló tanulási algoritmusok és keretrendszerek összessége, és nagyjából három ágra osztható:

1. *Felügyelet nélküli gépi tanulás*: az adatok particionálására és dimenziójának csökkentésére szolgáló adatbányászati technika. A felügyelet nélküli tanulás kiegészíti és általánosítja az adatcsökkentés statisztikai megközelítéseit, például a főkomponens-elemzést. A pénzügyekben alkalmazott nem felügyelt tanulás egyik példája a portfólióválasztáshoz használt K-means klaszterezés;
2. *Szupervizált gépi tanulás*: a regresszorok és regresszánsok közötti kapcsolat megtanulásának parametrikus vagy nem parametrikus, algoritmikus vagy prob-abilisztikus módszere. A felügyelt gépi tanulás általánosítja az olyan statisztikai technikákat, mint az OLS (ordinary least squares) re-gresszió vagy az idősoros módszerek, például az auto-regresszív modellek. A felügyelt gépi tanulás feltételezi, hogy a modell által hozott döntések nem befolyásolják a bemenetet. A felügyelt gépi tanulásnak csak néhány fajtája alkalmas nem stacionárius adatokra; és
3. *Erősítéssel tanulás*: a sztochasztikus irányítás egy olyan, visszacsatolással működő módszere, amely a bemenetek állapotát megváltoztató döntések alapján tanulja meg a politikát. A megerősítéssel tanulás sztochasztikus dinamikus programozást generál, és valószínűleg a kereskedés és a befektetés-menedzsment számára lesz a legnagyobb hatással. Ugyanakkor bonyolultsága miatt ez a legkevésbé kihasznált módszer a pénzügyekben. A pénzügyekben például a derivatívák árazása, az optimális fedezeti ügyletek, a Merton-féle portfólióprobléma és az optimális kereskedés végrehajtása.

A felügyelt és a felügyelet nélküli tanulás közötti különbségtétel általában jól érthető - ha az adatokat nem címkézzük fel egy válasszal, például a hitel státuszával, akkor a felügyelet nélküli tanulás a megfelelő megközelítés a minták és az adatok alacsonyabb dimenziós reprezentációinak felfedezésére. Ezzel szemben a címke jelenléte azt sugallja, hogy a felügyelt tanulás megfelelőbb. Ez a paradigma szigorúan a bemeneti adatok és a kimenet közötti térkép paramétereinek becslésére szolgál, a hiba képzési mintákon való minimalizálásával, azaz a kovariánsok feletti regresszió illesztésével. A teljesítmény általánosítása a regularizációs paraméterek becslésével érhető el a keresztellenőrzési adatokon.

Ha a paramétereket egyszer megtanultuk, akkor azokat nem frissítjük az új adatokra reagálva. Emiatt a felügyelt tanulás a tanulás "offline" formájának tekinthető, azaz a modell offline kalibrálása offline történik. Megjegyezzük, hogy kerüljük a modell statikusnak való megjelölését, mivel bizonyos típusú architektúrák esetén lehetséges olyan dinamikus modell létrehozása, amelyben a bemenet és a kimenet közötti térkép idővel változik. Például egy LSTM (Long Short Term Memory) vagy egy GRU (Gated Recurrent Unit) a Markov-kapcsolási modellekhez hasonlóan rejtett állapotváltozókat tart fenn, amelyek a bemenet-kimenet térképének idővel eltérő formáját eredményezik (Alonso et al., 2018).

A *felügyelt tanulás* során a "tanár" a gyakorlóhalmaz minden egyes adatpontjához pontosan a megfelelő kimenetet adja meg. Ez tekinthető a tanár "visszajelzésének", ami a felügyelt tanulás esetében azt jelenti, hogy az ágens minden egyes alkalommal, amikor az ágens egy új adatpontot osztályoz a képzési adathalmazban, a helyes címkével tájékoztatja az ágenst. Megjegyzendő, hogy ez ellentétes a felügyelet nélküli tanulással, ahol nincs tanár, aki helyes válaszokat adna az ML-algoritmusnak, ami úgy tekinthető, mint egy olyan környezet, ahol nincs tanár, illetve nincs visszajelzés a tanártól.

A megerősítéssel tanulás egy ágens segítségével döntések (*akciók*) sorozatát hozza meg a bemenet (*állapot-tér*) függvényében. A legfontosabb különbség ebben a környezetben a felügyelt tanuláshoz képest, hogy a tanár visszajelzése valamennyire a két véglet, a nem felügyelt tanulás (egyáltalán nincs visszajelzés) és a felügyelt tanulás között helyezkedik el, amely a megfelelő címkék megadásával visszajelzésnek tekinthető. Ehelyett az ilyen részleges visszajelzést a "jutalmak" biztosítják, amelyek a kívánt viselkedésre ösztönöznek, de anélkül, hogy kifejezetten utasítanak az ágensre, hogy pontosan milyen döntést kell hoznia, mint a felügyelt tanulásban.

A felügyelt tanulást jellemzően olyan alkalmazásokban használják, amelyek hagyományosan az OLS vagy idősoros regresszióra támaszkodnak, mint például az alapvető tényezők modellezése és az autoregresszív modellezés. A feedforward és a rekurrens neurális hálózatok közvetlenül általánosítják a lineáris regressziót nemlineáris regresszióra, illetve az autoregresszív modelleket nemlineáris autoregresszív modellekre. Jól skálázódnak a nagy dimenziós bemeneti adatokhoz, és például az OLS-ben elkerülhető a kolinearitási probléma. Emellett inkább a mintán kívüli teljesítményre, mint a mintán belüli teljesítményre illeszkednek.

A megerősítéssel tanulás a legmegfelelőbb paradigma, amikor egyrészt egy hasznossági függvényt kell optimalizálni, másrészt a döntések megváltoztatják az állapotot. Például egy kereskedési végrehajtási algoritmusnak maximalizálnia kell a kockázattal korrigált hozamot, és csökkentenie kell az árhatás hatását azáltal, hogy a nagy blokkos megbízásokat kisebb megbízásokra osztja fel. A pénzügyi modellezésben a megerősített tanulás célja gyakran az a probléma, hogy pontosan hogyan kell végrehajtani, likvidálni vagy fogyasztani például a kockázattal korrigált hozam optimalizálása érdekében.

2.1 Szükséges-e a mélytanulás a pénzügyekben?

A mélytanulás egyfajta felügyelt tanulási módszer, amely neurális hálózati rétegek összeállítását foglalja magában. További rétegek hozzáadásával általánosítja a lineáris regressziót vagy a logisztikus lineáris regressziót - lehetővé téve a bemenet és a kimenet közötti bármilyen típusú nem lineáris függvény ábrázolását. Valójában egy egy rejtett réteggel rendelkező, úgynevezett "előrecsatolt" neurális hálózat (más néven sekély tanuló) rendelkezik egy univerzális reprezentációs tétellel¹, amely bizonyítja, hogy a hálózat bármilyen Borel-féle mérhető függvény közelítésére használható a bemenet és a kimenet között. Miért van tehát szükségünk a neurális hálózatok bonyolultabb típusaira, beleértve a mélytanulást is?

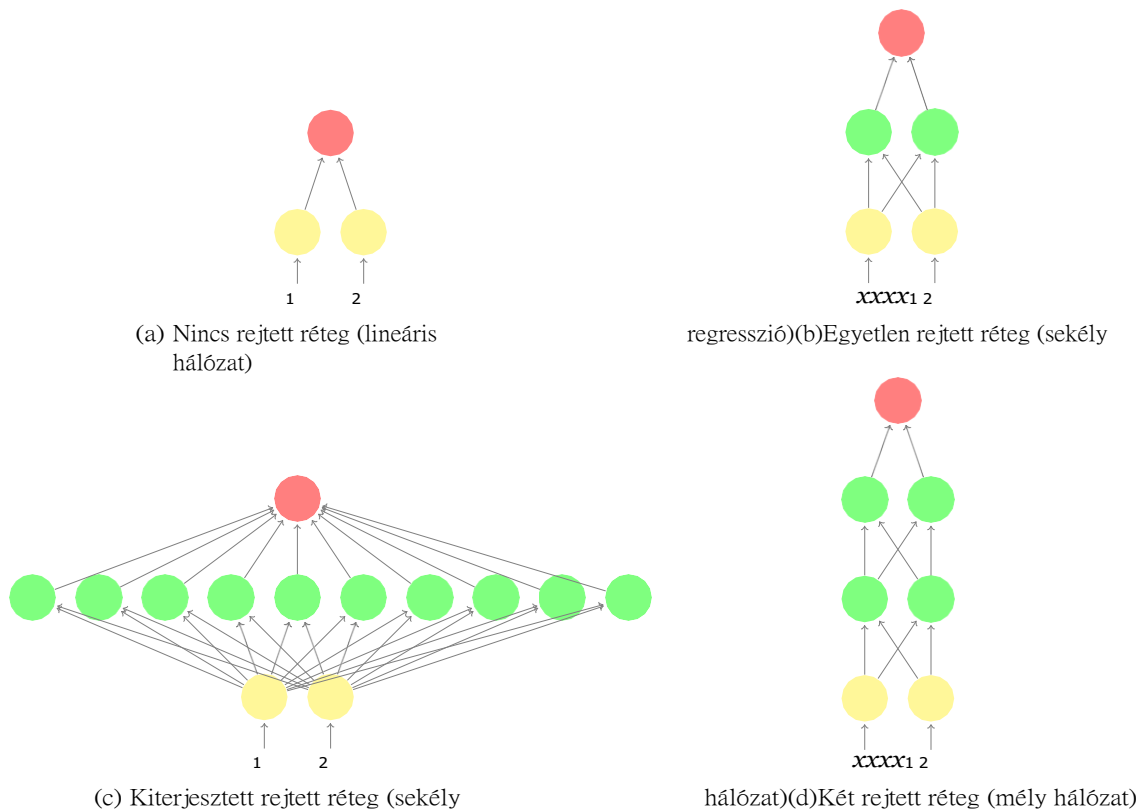
A természetes nyelvi feldolgozáson vagy a szövegbányászaton kívül kiderült, hogy a legtöbb pénzügyi felhasználási esetben a mélytanulás nem feltétlenül szükséges, hanem inkább kényelmi funkció, amely lehetővé teszi a különböző gráfok visszaterjedésen keresztüli összekapcsolását. Bizonyos alkalmazásokban, különösen nagy dimenziós bemenetek és erős nemlinearitás esetén előnyös lehet a mély tanulás használata egy "fanatikus" sekély hálózattal szemben (Sirignano et al., 2016; Dixon et al., 2016; Gu et al., 2018). Az előny azonban elsősorban a hatékonyságban rejlik - kevesebb paraméterre lehet szükség, ami gyorsabb számításokat eredményezhet². A mély hálózatok valójában értelmezhetőek³ - a kimenet érzékenysége az egyes bemeneti változókra meghatározható, így az az elképzelés, hogy fekete dobozok lennének, túlságosan leegyszerűsítő (Dixon és Polson, 2019). A mélytanulás a neurális hálózatok szinonimájává vált, és a neurális hálózatok képzésére szolgáló szabványos eszközöket, például a PyTorch-ot és a TensorFlow-t mély hálózatokra tervezték.

A valóságban a legtöbb neurális hálózatnak, amelyet a pénzügyekben használnak, nem kell mélynek lennie, és a hype

¹Az univerzális reprezentációs tétel szerint csak egy rejtett rétegre van szükségünk (Kolmogorov, 1957; Arnold, 1957).

²A mély hálózatok implicit módon önszabályozóak (Martin és Mahoney, 2018), ami azt jelenti, hogy kevesebb paraméterre van szükség, mint egy sekély hálózat esetében, hogy ugyanazon adatokon a mintán kívüli teljesítményt érjük el.

³Hozzá kell tennünk azt a figyelmeztetést, hogy a súlyok és torzítások nem hordoznak egyéni jelentést, mint mondjuk a lineáris regresszióban, hanem inkább algebrai kifejezések írhatók a jakobiánus és a kölcsönhatásokhoz, amelyek felhasználhatók a jellemző fontosságának rangsorolására.

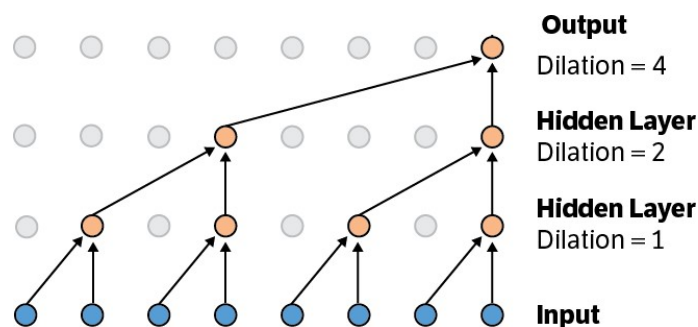


1. ábra: Egyszerű kétváltozós előrecsatolt hálózatok rejtett réteggel és anélkül. A sárga csomópontok a bemeneti változókat, a zöld csomópontok a rejtett egységeket, a piros csomópontok pedig a kimeneteket jelölik. A rejtett rétegek nélküli feedforward hálózat egy lineáris regresszor. Az egy rejtett réteggel rendelkező előrecsatolt hálózat sekély tanuló, a két vagy több rejtett réteggel rendelkező előrecsatolt hálózat pedig mély tanuló.

nem a valóságon alapul. Tisztázzuk, hogy bár az a felfogás, hogy a mélytanulás a neurális hálózatok szinonimája, megközelítőleg érvényes, nem igaz, hogy a mélytanulás a mesterséges intelligencia szinonimája. A mesterséges intelligencia az algoritmusok és módszerek sokkal nagyobb univerzumát jelenti, amelyek közül soknak még nem sikerült debütálnia a pénzügyekben.

A mélytanulás használata a piaci adatokból történő pénzügyi modellezéshez azonban nem csak negatív. Mint kiderült, van jó oldala is. Az LSTM-eken és a GRU-kon kívül a pénzügyi idősorok előrejelzése szempontjából igen relevánsnak tűnő mélytanulási jelöltet még nem sikerült határozottan kipróbálni a pénzügyi adatokon. A neurális hálózatok egy bizonyos típusa, az úgynevezett konvolúciós neurális hálózat, pénzügyi idősorokra alkalmazva képes különböző időskálájú minták megragadására (Borovykh et al., 2017; van den Oord et al., 2016). A *dilatált konvolúciós neurális hálózatnak* nevezett architektúra képes megtanulni az autokorrelációs struktúrákat különböző frekvenciákon, például napi, heti, havi, éves történelmi mintákat. Minden réteg más-más időskálát ragad meg, hasonlóan a különböző frekvenciájú autoregresszív modellek összeállításához.

A pénzügyi gépi tanulás területe szenvedett attól, hogy túl könnyen beleszeretett az inno-



2. ábra: Három rétegből álló tágított konvolúciós neurális hálózat. A receptív mező a következő $r = 8$, azaz egy kimeneti értéket nyolc bemeneti neuron befolyásol. Forrás van den Oord et al. (2016).

vations from AI research - a mélytanulás az egyik példa arra, hogy a mesterséges intelligencia kutatóitól közvetlenül kölcsönvett ötletek egyszerűen szögletes csapok kerek lyukakba ültetéséhez vezettek. Véleményünk szerint széles körű tudatosságra van szükség az emberi tényezőkkel kapcsolatban, amelyek ezekhez az átvételi problémákhoz vezetnek, és egy keretrendszerre, vagy legalábbis iránymutatásokra van szükség arra vonatkozóan, hogy miként lehet a mesterséges intelligenciával kapcsolatos kutatási innovációkat a pénzügyi modellezés területére átültetni. A cikk hátralévő részében ezt kívánjuk tovább vizsgálni.

3 A gépi tanulás négy lovasa a pénzügyekben

Annak ellenére, hogy a felügyelt gépi tanulás (Dixon et al., 2016; Heaton, Polson és Witte, Heaton et al.; Bayer és Stemper, 2018; Feng et al., 2018) és a megerősítéses tanulás (Halperin, 2017; ~~It~~ et al., 2018) pénzügyi modellezésben való sikeres alkalmazásainak lenyűgöző sora ellenére ugyanolyan, ha nem nagyobb tömegű anekdotikus történetek vannak a gépi tanulásról, amelyek az alapok kudarcához, a pénzügyi adatok gyenge teljesítményéhez vagy akár a magvető kutatási projektek meghiúsulásához vezetnek. Míg a természetes nyelvi feldolgozást és szövegbányászatot igénylő problématerületek nagymértékben a jól bevált mérnöki gyakorlatok közé tartoznak, a pénzügyi modellezés a mérnöki és a pénzügyi modellezési tapasztalatok keverékét igényli, ami még mindig ismeretlen terület. Tapasztalataink és az iparági szakértőkkel folytatott ~~komunikációk~~ alapján elemezzük a gépi tanulás pénzügyi intézményekben történő bevezetését nehezítő néhány gyakori problémát, és bemutatjuk a következő négy lovasat:

- *A barkácsoló*: A svájci bicskával felfegyverkezve a barkácsoló statisztikailag analfabéta, aki képtelen és nem hajlandó tudományágát a pénzügyi idősorelemzéssel, a pénzügyi modellezéssel és ökonometriával, valamint a dinamikus programozással integrálni. Úgy tűnik, hogy ez a képtelenség részben abból ered, hogy a gépi tanulás a pénzügyeken kívüli, gyakran a Kaggle-en keresztül elérhető egyszerű mintafelismerési problémákra való alkalmazásában szokott sikereket elérni. A végeredmény egy olyan Rube-Goldberg-gép, amely nem képes védhető, értelmezhető eredményeket produkálni, amelyeket össze lehet egyeztetni egy jobban megalapozott módszertannal. Ráadásul az érthetetlen hiperparaméterek megváltoztatása jelentősen megváltoztatja az eredményt, és a legjobb esetben is heurisztikákat használnak az eredmény stabilizálására.
- *A történész*: A történész azt állítja, hogy a világ továbbra is olyan marad, mint amilyen a múltban volt. Előrejelzések készítésekor gyakori hiba, hogy a kovariancia stacionaritását feltételezik.

az adatok. Az idősor-előrejelzési módszerek, például a rekurrens neurális hálózatok csak akkor használhatók megbízhatóan, ha az adatok autokorreláltak és a kovariancia stacionárius.⁴ A nem stacionárius adatokon szinte minden felügyelt gépi tanulási módszerrel történő magas előrejelzési teljesítmény elérésére tett kísérlet valószínűleg hiábavaló. Nem stacionárius adatokra csak a Gated Recurrent Units (vagy a szorosan kapcsolódó LSTM-ek) és a változási pontok detektálásán alapuló gépi tanulási módszerek alkalmasak.

- *A bányász* : A bányász rövidlátóan abból a feltevésből indul ki, hogy egy lépésre van attól, hogy aranyat találjon, anélkül, hogy figyelembe venné a tágabb összefüggéseket, hogy a piacon más résztvevőkkel versenyezzen. Miután a modell viszonylag nagy pontosságú jelet generál egy algó-kereskedési stratégiához, aztán tesztelik a piaccal szemben, vagy visszatesztelik, de csak középserű eredményt hoznak. A bányász ezután kimerítően alkalmazza a rendelkezésére álló eszközarzenáljában rendelkezésre álló összes gépi tanulási osztályozó vagy regressziós technikát, mielőtt végül feladja, és vonakodva megállapítja, hogy a stratégia teljesítménye a felügyelt gépi tanulás segítségével nem teljesített jobban, mint a lineáris regresszió. Az algoritmikus kereskedés szakértői jól tudják, hogy a késleltetés, a csúszás, az árhatás és a kereskedés végrehajtása fontos szerepet játszik a stratégia teljesítményében, és a megerősítő tanulás a megfelelő paradigma e hatások figyelembevételére. Általánosságban elmondható, hogy a nagy pontosságú felügyelt tanulási mintafelismerési módszerek nem feltétlenül eredménymaximalizáló jeleket eredményeznek.
- *A puritán*: Az eredményeknek meg kell felelniük az elvárásaiknak és hiteiknek, mielőtt hajlandóak lennének támogatni a gépi tanulást, és csak akkor, ha az automatizálás eszközeként használják. A puritán képtelen elfogadni, hogy a gépi tanulás új és értelmes mintákat fedezhet fel, amelyek nincsenek összhangban a jól bevált elméletekkel és a pénzügyi kánonjaiba beírt stilizált tényekkel. Ezzel figyelmen kívül hagyják a gépi tanulásnak mint a kiugró értékek és nem lineáris hatások megragadására szolgáló eszköznek néhány legfontosabb előnyét.

Minden lovas egy kóros gondolkodásmódot testesít meg, ha a végletekig elviszik. Természetesen az innováció és a kockázatkezelés gondolkodásmódja szöges ellentétben állhat egymással, és a kísérletezés és az egészséges szkepticizmus egy bizonyos fokozata a kutatás velejárója. Tapasztalataink szerint néhány egyszerű iránymutatás kikövezheti az utat a gépi tanulás pénzügyi modellezésben való sikeresebb alkalmazása felé.

4 A hatékony gépi tanulás alapelvei a pénzügyekben

Azzal a nagy fenntartással, hogy nem helyettesíthető a kritikus emberi gondolkodási készség, a gépi tanulás elméletének ismerete, az alkalmazási terület szakértelme, elegendő és megbízható adat⁵ és technikai erőforrás az adatok feldolgozásához és a gépi tanulás alkalmazásához, most néhány iránymutatást ismertetünk a gépi tanulás sikerének javítására a befektetéskezelésben:

1. *A probléma meghatározása*: A megoldandó probléma meghatározása. Gondosan mérlegelje, hogy a gépi tanulás alkalmazása egyértelműen az adatszimatoláson alapul-e - a kereskedési stratégiát és a végrehajtást külön kezelve -, vagy a problémának van-e ütemezési összetevője, és megerősítő tanulást igényel-e. Határozza meg továbbá, hogy számítási szempontból megvalósítható-e a teljes adathalmaz elemzése, vagy először adatcsökkentésre (felügyelet nélküli tanulásra) van szükség.

⁴Az egyetlen eset, amikor ez nem igaz, az, ha a modellező hajlandó a modellt gyakran újratréningelni, ami nem

biztos, hogy praktikus, és stabilitási problémákat eredményez.

⁵Az adatoknak kellően magas jel-zaj aránnyal kell rendelkezniük.

2. **Modellezési feltételezések:** A modellezési megközelítés fő feltételezései: A machine learning módszer feltételezi, hogy a megfigyelések azonosak és függetlenül oszlanak meg? Ha az adatok idősorok, a modell feltételez-e kovariancia-stacionaritást? Alkalmazza a táblázatban felsorolt alapvető statisztikai tesztek az adatok jellemzésére.
3. **Fejlessze az intuíciót:** Állítson fel egy játékmotelt, a teljes probléma egy részhalmazát, hogy elsajátítsa a gépi tanulás kényelmét egy olyan környezetben, ahol kevesebbet lehet elrontani. A neurális hálózatok esetében például a modell lineáris változatával hozhatunk létre egy alapszintet, azaz amikor a neurális hálózat funkcionálisan egyenértékű a lineáris regresszióval.⁶;
4. **Védhető eredmények:** Tervezze meg a kísérletet úgy, hogy a modell kimenete megmagyarázható legyen, és abszurdum redukcióval védhető legyen. Ez nem azt jelenti, hogy a modellnek statisztikailag értelmezhetőnek kell lennie, hanem azt, hogy az esetleges téves osztályozások vagy nagy hibák elszámolhatók legyenek. Sok esetben szükség lehet a játékmotellhez való visszatérésre a rejtély kibontásához.
5. **Diagnosztika:** Többféle diagnosztika alkalmazása, mind a gépi tanulásból, mind a statisztikából, az adatok és a torzítás-változás, valamint a modell szignifikanciájának jellemzésére egy alapvonalhoz képest. Az A. függelék tartalmaz egy rövid összefoglalót a pénzügyekben a gépi tanuláshoz hasznos néhány legfontosabb statisztikai tesztől;
6. **Tartsa egyszerűnek:** Minimalizálja a bonyolult algoritmusok és widgetek használatát a minél kevesebb paraméterre támaszkodó és a lehető legátláthatóbb megközelítések javára;
7. **A hasznossági függvény megválasztása:** A konkrét algoritmusok teljesítményének elemzése előtt vizsgálja felül a gépi tanulási algoritmusok képzéséhez használt veszteségfüggvényre vonatkozó feltételezés megfelelőségét. Megfelel-e például a közönséges legkisebb négyzetek veszteségfüggvénye, amikor elemzők vállalati nyereség-előrejelzéseit elemzi?;
8. **Megoldási korlátok:** Elemezze az összes lehetséges előzetes nézetet vagy korlátozást, amelynek a várható megoldásnak meg kell felelnie (pl. pozitív ár, ritkaság, konvexitás, arbitrázsmentesség stb.), és ezeket az előzetes nézeteket kényszerítse ki a megoldásra, akár kemény, akár puha korlátozások formájában (regularizáció);
9. **Elfogultak az adatai?** Sok esetben a gépi tanulási algoritmusok képzéséhez rendelkezésre álló adathalmazok nem kiegyensúlyozottak. Például éppen az az esemény, ami érdekel, egy feketerigó. Az adatok rossz egyensúlya az Ön előrejelző modelljének torzításaiba terjedhet, amit a statisztikában "ellenőrzési paradoxonként" ismerünk. Az algoritmus tervezett alkalmazásától függően az ilyen hatások potenciálisan károsak lehetnek vagy nem - ha a jövőbeli adatok várhatóan megtartják a képzési adatok félreegyensúlyát, akkor figyelmen kívül hagyhatja, de egyébként korrekciókat kell végrehajtania;
10. **Jellemző mérnöki munka:** Ha lehetséges, kerülje a kísértést az idő előtti funkciómérésre, a különböző gépi tanulási modellek közötti váltásra anélkül, hogy előbb a jelet visszatesztelési vagy élő szimulációs környezetben értékelné. Különösen a stratégia teljesítménytartományát határozza meg, felmérve, hogy mi a legjobb és a legrosszabb eset P&L értéke a tökéletes jel (vagy a mintán belüli jel), illetve a fehér zaj használatával. Jellemezze, hogy a stratégia hogyan veszít és hogyan keres pénzt az alábbiak alapján

⁶A neurális hálózatok egyik előnye, hogy az aktiválási függvények kikapcsolhatók, így a hálózat funkcionálisan lineáris regresszióvá válik, ami könnyebben érthető. Ebben az esetben van egy potenciális degeneráció; létezhetnek "lapos irányok" - hiperfelületek a paraméterterben, amelyek pontosan ugyanolyan veszteségfüggvényrel rendelkeznek.

jelet, mielőtt értékes időt fordítanánk a funkciótervezésre. Ha a jel rosszul van beállítva a stratégiához és a piachoz, nincs garancia arra, hogy a funkciótervezés javítani fogja a stratégia teljesítményét.

5 Példák a gépi tanulás legjobb gyakorlataira a pénzügyekben

Bár a pénzügyi modellezésben a gépi tanulásnak számtalan példája van, mi csak néhány archetipikus problémára összpontosítunk, amelyek mindegyike különböző, de fontos módon illusztrálja az előző szakaszban a legjobb gyakorlatra vonatkozó iránymutatásokat.

5.1 Tényezőmodellezés mélytanulással

Lineáris keresztmetszeti faktor modellek, mint például (Fama és MacBeth, 1973), FF(3) és FF(5) (Fama és French, 1993, 1992, 2015) és a BARRA-faktor modellek (lásd Rosenberg és Marathe (1976); Carvalho et al. (2012)) egyszerűségük és gazdasági ~~értékük~~ miatt vonzóak, kereskedhető portfóliókat generálnak. A BARRA-modellben a faktormegvalósulások becslése általánosított legkisebb négyzetek regressziójával történik.

A legkisebb négyzetek lineáris regressziója azonban rosszul kifejezhető, és független Gauss-hibákra támaszkodik, ami a *nemlineáris* modellek mérlegelését indokolja. A nemlinearitások bevezetése és a kölcsönhatási hatások beépítése azonban nehezebb feladat lehet. A vagyongazdálkodók új, előrejelző vállalati jellemzőket keresnek az olyan anomáliák magyarázatára, amelyeket a klasszikus tőkeeszköz-árazási és faktormodellek nem ragadnak meg. A közelmúltban számos független, adattudományi megközelítésben gyökerező empirikus tanulmány rámutatott a vállalati jellemzőkkel és más közös tényezőkkel kapcsolatos, gazdaságilag értelmezhető prediktorok nagyobb számának használatának fontosságára Moritz és Zimmermann (2016); Harvey et al. (2015); Gu et al. (2018); Feng et al. (2018). Gu et al. (2018) egy több mint 30 000 egyedi részvényt tartalmazó adathalmazt elemeznek egy 60 éves, 1957 és 2016 közötti időszakra vonatkozóan, és több mint 900 alapjelet határoznak meg. Mind Moritz és Zimmermann (2016), mind Gu et al. (2018) kiemeli az OLS-regresszió elégtelenségét a változéválasztás során a nagy dimenziójú adathalmazok esetében.

A szakirodalomban számos olyan újabb megjelent cikk található, amely bizonyítékot talál a nemlineáris regressziós technikák jobb teljesítményére az alapvető faktormodellek, például a regressziós fák és a neurális hálózatok esetében. (Gu et al., 2018; Feng et al., 2018; Dixon és Polson, 2019). A mély versus sekély architektúra árnyalatai nem szabad, hogy háttérbe szorítsák a nemlineáris regresszió fontosságát. A nemlineáris regresszióhoz számos technika létezik, de a neurális hálózatok bizonyíthatóan minden nemlineáris függvényt képviselnek. Ugyanez nem igaz a fákra, kivéve, ha azok végtelenül mélyek.

A Dixon és Polson (2019) alábbi, alapvető tényezőkre vonatkozó neurális hálózatában bemutatjuk az ML bevezetésének ajánlott gyakorlatát. A kísérletek részletei nem teljesek, és vannak más tesztek és kísérletek, amelyeket el kell végezni, mint például a regularizáció mértékének hangolása és a tanulási görbék megtekintése, hogy megértsük a mintaméret hatását a torzítás-variáció kompromisszumra. Az eredményeket ráadásul csak egy egyszerű, hat alapvető tényezőtől álló adatsorra alkalmazzák - mindazonáltal a robusztus kísérleti tervezés elve itt is érvényes marad.

A probléma megfogalmazása: A BARRA megközelítést követjük - a feedforward hálózat bemenetei az egyes (eszköz, idő) párok havi jelentett tényezőterhelései, a kimenet pedig a következő

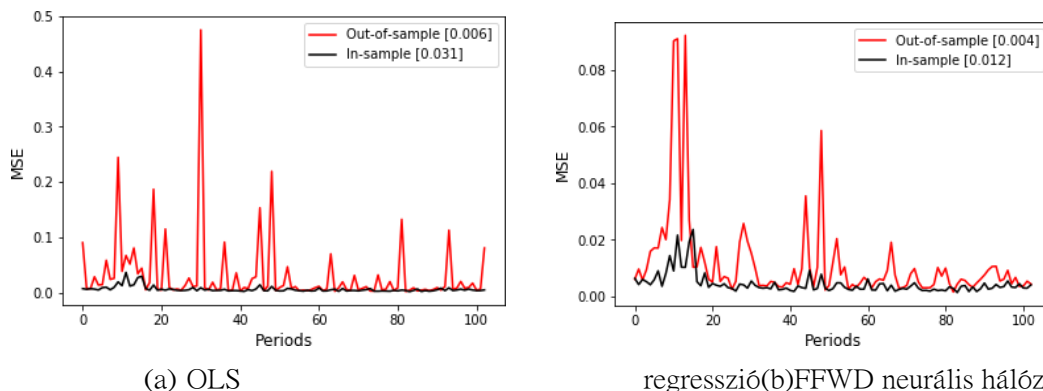
1 hónapos megtérülés⁷. A probléma az, hogy megtaláljuk a nem lineáris térképet a faktorterhelések és az eszközhozamok között a részvények univerzumában egy adott történelmi időszakra vonatkozóan. A lineáris regresszióban a modell érzékenysége az egyes faktorokra megfelelnek a faktorok realizációinak. Látni fogjuk, hogy ez a koncepció hogyan általánosítható a (sekély vagy mély) neurális hálózatokra is.

Pontosabban, neurális hálózatokat használunk arra, hogy megtaláljuk a faktorterhelések és az eszközhozamok közötti térképet. Ha K alapvető tényező és N eszköz van, akkor a bemeneti adat az ismert időinvariáns tényezőterhelések (béták): ahol $B = [\mathbf{1} \mid \beta_1 \mid \dots \mid \beta_K]$ az ismert tényezőterhelések (béták) $N \times K + 1$ mátrixa: $\beta_{i,k} := (\beta\mathbf{k})_i$ az i eszköz kitettsége a k tényezőnek t időpontban. A kimenet az N eszköz hozama, \mathbf{r}_t , t időpontban:

$$\mathbf{r}_t = F_t(\mathbf{B}) + \epsilon_t \quad (1)$$

A hálózatot minden egyes havi időszakra keresztmetszeti regressziós modellként képzik ki a kiválasztott univerzumban lévő összes eszközre - ebben az esetben az S&P 500 index 218 eszközére. Minden egyes múltbeli időszakra a neurális hálózatot újratranítjuk, és összehasonlítjuk a teljesítményt a realizált eszközhozamokkal.

Az alapszint megállapítása: A neurális hálózat erősségének felmérése érdekében bemutatjuk, hogyan állítjuk fel a neurális hálózat alapvonalát OLS-regresszióval, és összehasonlítjuk egy egyszerű részvénykiválasztási stratégia eredményét az ML jel és egy fehér zajjal segítségével. Megjegyezzük, hogy ha a neurális hálózat alulteljesítette az OLS-regressziót, akkor egyszerűen az aktiválási függvények kikapcsolásával (azaz lineáris aktiválás használatával) orvosoljuk a hibát, és ellenőrizzük, hogy az eredmények átlagosan hasonlóak-e.⁸



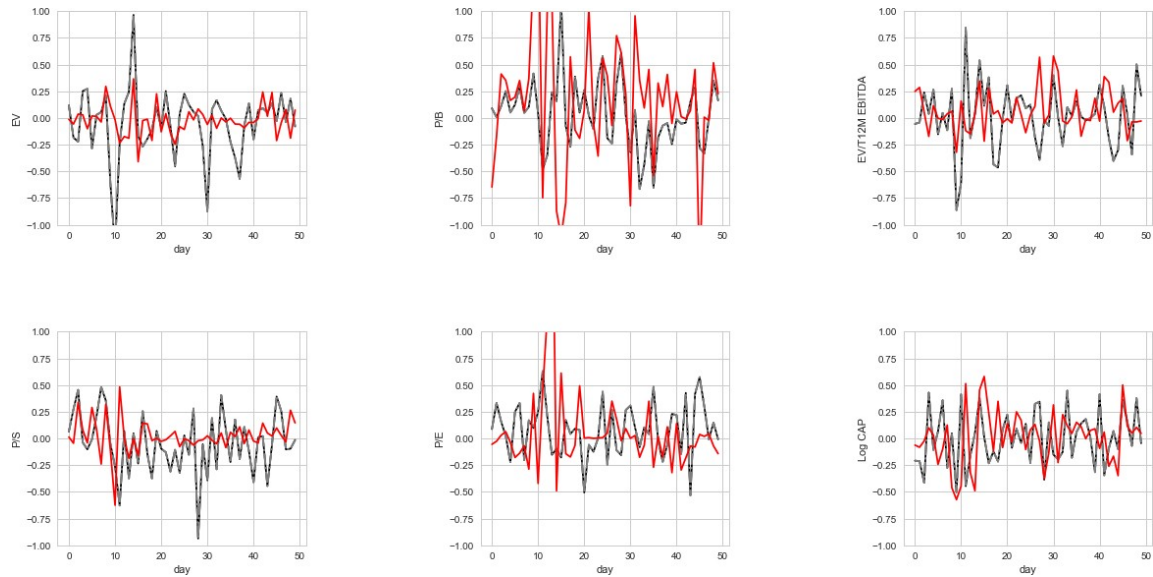
3. ábra: Ez az ábra az OLS becslő és a feedforward neurális hálózat mintán kívüli teljesítményét hasonlítja össze, az átlagos négyzetes hiba (MSE) alapján.

Értelmezhetőség: A neurális hálózatok valójában értelmezhetőek - a kimenet érzékenysége a bemenetekre és a kölcsönhatási feltételek analitikusan kiértékelhetőek. Az ábra az érzékenységeket mutatja⁴

⁷A 3 hónapos megtérülést a gyakorlatban gyakrabban használják.

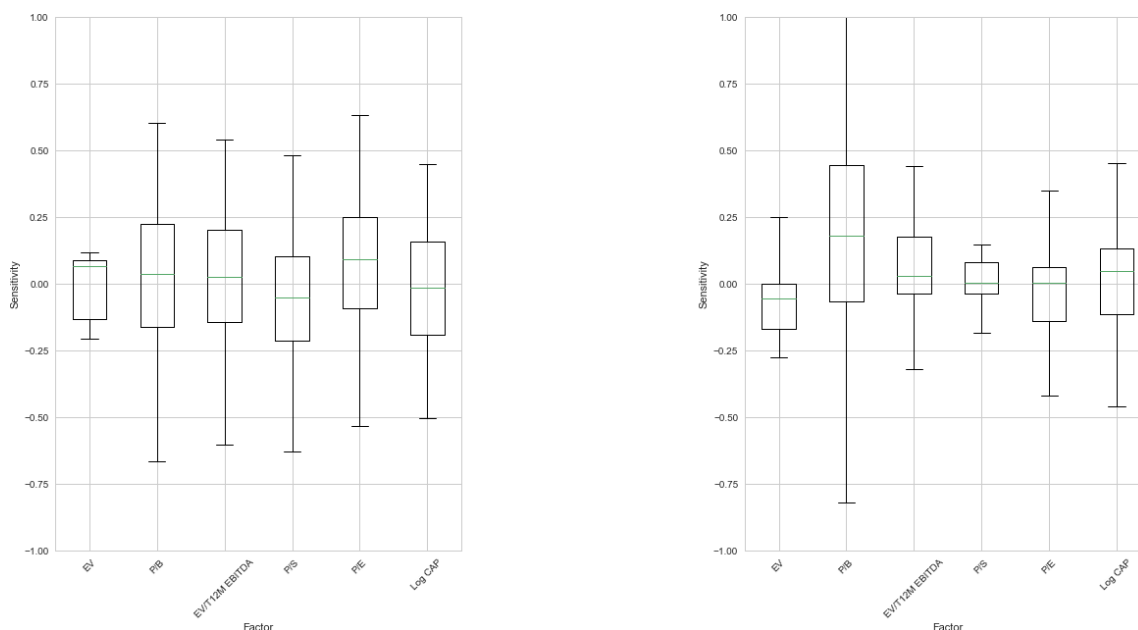
⁸Emlékezzünk vissza, hogy a rejtett egységek lineáris aktiválásával rendelkező neurális hálózat, legyen az sekély vagy mély, lineáris regresszió.

az OLS regressziós modell és a neurális hálózati modellek hat fundamentális tényező mindegyikére a 2014 novemberétől 2018 novemberéig tartó 48 hónapos időszakban. A fekete vonal a neurális hálózaton alapuló módszerünk által értékelt érzékenységeket mutatja. A piros vonal ugyanezeket az érzékenységeket mutatja a lineáris regresszió használatával.



4. ábra: Ez az ábra a modell érzékenységét mutatja az egyes tényezőkre a 2014 novemberétől 2018 novemberéig tartó 48 hónapos időszakban. A fekete vonal a neurális hálózaton alapuló módszerünk által értékelt érzékenységeket mutatja. A piros vonal ugyanezeket az érzékenységeket mutatja lineáris regresszióval. Megjegyzendő, hogy a tényezők közötti összehasonlítás megkönnyítése érdekében a grafikon skálázását úgy állítottuk be, hogy az y-tengely tartománya rögzített.

Az ábra az egyes tényezőkre vonatkozó érzékenységek eloszlását hasonlítja össze egy hónapos 48 időszakon keresztül novemberétől 2014 novemberéig a neurális hálózatok és az OLS-regresszió által 2018-ben becsült értékek szerint.

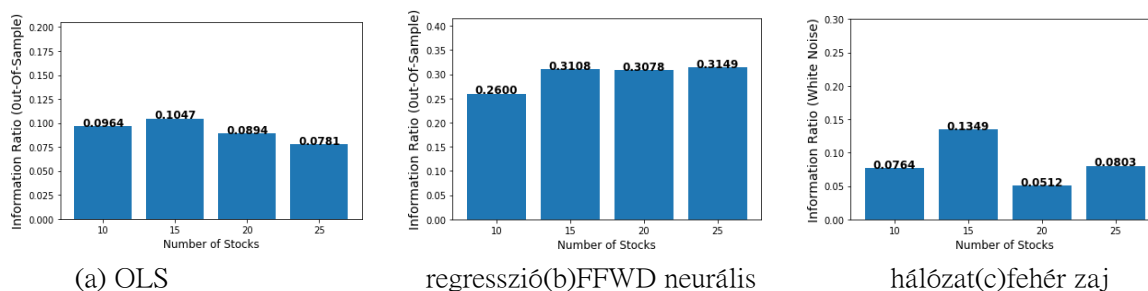


5. ábra: Az egyes tényezőkre vonatkozó érzékenységek eloszlása a 2014 novemberétől 2018 novemberéig tartó 48 hónapos időszakban a neurális hálózat által becsült értékek szerint (balra). Ugyanezek az érzékenységek általános lineáris regresszióval (jobbra). Megjegyzendő, hogy a tényezők közötti összehasonlítás megkönnyítése érdekében a grafikon skálázását úgy állítottuk be, hogy az y-tengely tartománya rögzített.

Stratégiai teljesítmény: Végezetül, bár nem olyan átfogó eredmény, amely következtetést indokolna, bizonyítékot szolgáltatunk arra, hogy neurális hálózati faktor modellünk magasabb információs arányokat generál, mint a lineáris faktor modell, amikor az univerzumunkból származó portfóliók rendezésére használjuk. A 6. ábra egy olyan portfólióválogató stratégia információs arányait mutatja, amely a legmagasabb előre jelzett havi hozammal rendelkező n részvényt választja ki. Az információs arányokat különböző méretű portfóliókra értékeljük, az S&P indexet500 használva viszonyítási alapként.

A megfigyelések szerint az információs arányok körülbelül háromszorosára nőnek. Kontrollként véletlenszerűen kiválasztott portfóliók is szerepelnek.⁹ Néhány esetben megfigyelhetjük, hogy bár a lineáris regresszióban a tényezők alig haladják meg a fehér zajt, a neurális hálózati modellben használva valóban előrejelzőek. Még a tényezők e korlátozott halmaza esetében is pozitív információs arányokat lehet produkálni a neurális hálózatokkal.

⁹Megjegyzendő, hogy az információs arányok azért mindig pozitívak, még véletlenszerű portfólióválasztás esetén is, mert csak olyan részvényekből álló univerzumot határoztunk meg, amelyekről elegendő historikus adat áll rendelkezésre a novemberétől 2014novemberig terjedő időszakra vonatkozóan. 2018.



6. ábra: Ez az ábra egy olyan portfólióválasztási stratégia információs arányait mutatja, amely az univerzumból a legmagasabb előre jelzett havi hozammal rendelkező n részvényt választja ki. Az információs arányokat különböző méretű portfóliókra értékeljük. Az információs arányok a mintán kívüli, előre megjósolt eszközhozamokon alapulnak (balra) általánosított lineáris regresszió; (középen) neurális hálózatok; és (jobbra) véletlenszerű, előrejelző modell nélküli.

Ebből az első példaalkalmazásból az a tanulság, hogy a helyes kísérlettervezés elvei ugyanolyan fontosak, mint a gépi tanulási módszer. Ha egyszerűen csak kipróbálunk egy kész neurális hálózatot, és arra a következtetésre jutunk, hogy az "nem működik", azt a mintán kívüli teljesítményre vonatkozó bizonyítékokkal kell alátámasztani. Mivel a rejtett egységek lineáris aktiválásával működő neurális hálózat OLS-regresszió, először egy standard OLS-regressziós eszközzel kell összehasonlítható eredményeket megállapítani. Csak ezután térjen vissza a nemlineáris aktivált egységekre, és hangolja az L regularizációt, hogy általánosítási teljesítményt kapjon a mintán kívüli előrejelzésekre. Általánosságban azt kell gyanítanunk, hogy a neurális hálózatok felülmúlják a lineáris regressziót, de ez csak elegendő képzési adat és a torzítás-variáció kompromisszum gondos kezelése esetén valószínűsíthető kereszt-validálással. Bár itt nem mutattuk be, a tanulási görbék segítségével megállapítható a képzési minta méretének növelésének hatása a torzítás-variáció kompromisszumra (lásd például Dixon (2018)).

Talán észrevették, hogy a fenti kísérletben a kísérleti eredményeknek volt egy időösszetevője. Szigorúan véve a faktorterhelések keresztmetszeti és időbeli (azaz paneladat) adatok. Az előrejelzési horizont azonban jellemzően negyedéves, és a töltesek havonta frissülnek. Emiatt arra számíthatunk, hogy a modellt minden hónapban újra kell tanítani, és a modell dinamikus aspektusa figyelmen kívül hagyható. Természetesen a késleltetett faktortöltés fontos lehet a mintán kívüli teljesítmény szempontjából, és hasznosak lehetnek az idősoros modellek, például a rekurrens neurális hálózatok.

5.2 Előrejelzés RNN-ekkel

Az eszközárak idősorokból történő előrejelzése hagyományosan a pénzügyi közgazdaságtan területe. Emlékezzünk arra, hogy mivel a sima RNN-ek vektoros auto-regresszív modelleket általánosítanak, használhatjuk őket előrejelzésre történeti inputokból, akár endogén, akár exogén inputokból. Továbbá, a dinamikus RNN-ek, mint például a GRU-k és az LSTM-ek, akkor használhatók, ha az adatok nem stacionáriusak, dinamikus auto-regresszív struktúrát biztosítva.

A probléma megfogalmazása: A probléma az, hogy hogyan lehet megjósolni az egy lépéssel előremenő középárakat a Bitcoin történelmi középfolyamainak egyenlő távolságra lévő időbélyegeiből. Természetesen számos oka van annak, hogy alternatív bemeneti adatokat használjunk, különösen nagyobb gyakorisággal, de ennek a feladatnak az a célja, hogy

az idősoros adatokból gépi tanulással történő előrejelzés főbb kihívásainak és iránymutatásainak bemutatása.

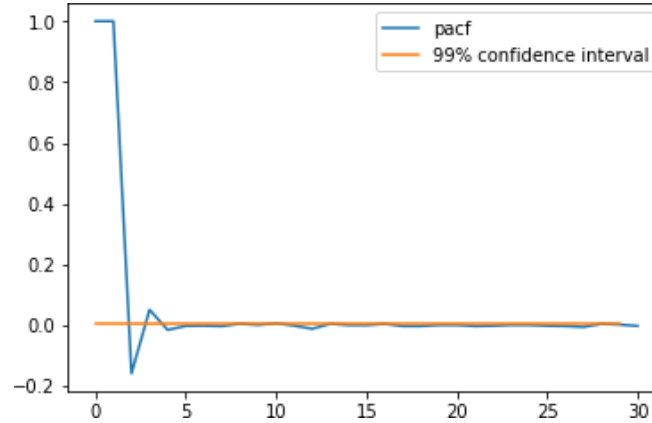
Statisztikai tesztek: Az RNN-ek egyváltozós idősorok modellezésére történő alkalmazásához először is meg kell állapítanunk, hogy az idősor autokovariancia-stacionárius-e. A Bitcoin USD-ben kifejezett középárfolyamok 1 perces pillanatfelvételeiből vett 199918 mintát használva azt találjuk, hogy a kiterjesztett Dickey-Fuller (ADF) tesztstatisztika 2.0945, a százalékos 10^{th} érték felett (2.5667) van, p-értéke pedig .5667. Mivel 0.26 a tesztstatisztika arra a nullhipotézisre vonatkozik, hogy az adatok nem stacionáriusak, nem tudjuk elutasítani. Ebben az esetben a teszt magában foglalja a konstans és a trend kifejezés hatását.

Néha az idősor átalakítható a különbség vagy a detrending segítségével. Bár ennek a transzformációnak vannak előnyei, messze nem teljesen biztos, és a prediktív jel elvesztéséhez vezethet. Itt egyszerűen az eredeti idősoron folytatjuk az előrejelzést, de figyelmeztetjük az olvasót, hogy vizsgálja meg a differenciálást vagy a detrendinget.

Modell azonosítása: A gépi tanulás a pénzügyekben nem váltotta fel a klasszikus Box-Jenkins megközelítést. Inkább arra épül, és (i) modell-azonosítással; (ii) illesztéssel; és (iii) diagnosztikával jut el a "legjobb" modellhez, ahol a mintavételen kívüli teljesítmény lesz a siker végső mércéje. A modell azonosítása kihívást jelenthet az idősorok esetében -- vagy összehasonlíthatjuk a veszteséget a változó maximális késleltetések ("szekvenciahossz") között egy RNN-ben, mivel ez az AR-modell egy speciális típusa, amelynek paramétereinek száma (azaz a súlyok és torzítások) nem nő a szekvenciahosszal.

Ez a megközelítés, bár teljes mértékben bizonyított, meglehetősen számításigényes lehet, és a becsült részleges autokorrelációs függvényt (PACF) használhatjuk a modellbe bevonandó késleltetések számának meghatározására.

Ábra 7 mutatja a becsült részleges korrelációs függvényt (PACF) a Bitcoin középárfolyamok percenkénti pillanatfelvételeire. Egy egyszerű RNN jellemzője, hogy nulla PACF-et mutat, amint a késleltetés meghaladja a szekvencia hosszát. Más szóval, van egy cut-off p késleltetés után, ha a szekvencia hossza p . Azzal a fenntartással, hogy az adatok nem stacionáriusak, így a PACF időben változik, a becsült PACF-et a minta átlagos modellrendjének azonosítására tekintjük. A grafikon alapján a kezdeti modellképzést vagy 4 vagy késleltetéssel kezdenénk Ez 5. előnyösebb, mint a szekvenciahossz kitalálása, vagy egy újabb hiperparaméter hozzáadása a hangolási fázishoz, és elkerülhető a túlzott szekvenciahossz megadása, ami hosszabb képzési időt eredményez.



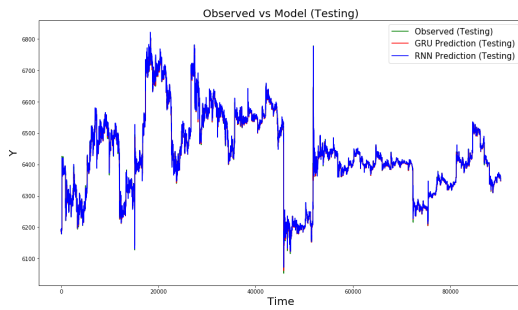
7. ábra: Ez az ábra a becsült parciális korrelációs függvényt (PACF) mutatja a percenkénti pillanatfelvételek Bitcoin-középfolyamokra.

Eredmények: A 8. a) ábrán látható 10 000 mintából¹⁰ álló tesztmintán végezzük el a következő perc előrejelzését egy RNN és egy GRU segítségével. Megjegyezzük, hogy a modell nem kerül újratanításra a tesztidőszak alatt, csak az ezt a tesztidőszakot közvetlenül megelőző képzési adatokat használja fel a modell a képzéshez.

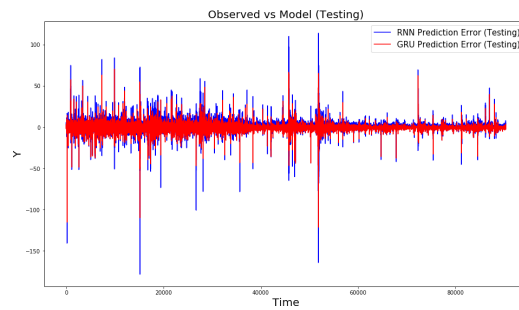
A várakozásoknak megfelelően a mintán kívüli hiba alacsonyabb a GRU esetében (lásd a 8. ábra b) pontját), amely nem feltételezi az adatok stacionaritását. Bár a hibák jelentéktelennek tűnhetnek, felhívjuk a figyelmet arra, hogy az y-skála USD-ben van. Alkalmazzuk a Mariano-Diebold-tesztet is, hogy megállapítsuk, hogy a két modell közötti különbség 99%-os szinten statisztikailag szignifikáns.

Bár a fenti kísérlet részletei ismét nem teljesek, a kísérlet és a modelltervezéshez szükséges statisztikai tesztek alkalmazásának fontosságát szemlélteti. További statisztikai tesztek is el kell végezni, például a Ljung-Box-tesztet a mintán kívüli reziduumokon annak megállapítására, hogy a modell alulilleszkedik-e azáltal, hogy nem ragadja meg az idősor teljes szerkezetét.

¹⁰Az előtekintés torzításának elkerülése érdekében a tesztelési kísérletet a gyakorlókérdés jövőjébe helyezzük.



(a) Egy lépés előre történő előrejelzés RNN-ekkel



(b) Egy lépés előre történő előrejelzési hiba

8. ábra: (balra) Ez az ábra összehasonlítja a következő perces sima RNN és GRU előrejelzéseket a Bitcoin10,000 középárfolyamok tesztmintájával. (Jobbra) A várakozásoknak megfelelően a mintán kívüli teljesítmény a GRU esetében alacsonyabb, mint a sima RNN esetében - az utóbbi stacionaritást feltételez.

Ezt a példát itt nem fejlesztjük tovább, ahogyan az előző példánál tettük, egy kereskedési stratégia backtesztelésével, amely a modell előrejelzését jelként használja. De ez természetesen szükséges lépés lenne az alfa generálására szolgáló előrejelzési modell előállításához. Leggyakrabban a kereskedés végrehajtásával kapcsolatos aggályok, mint például a csúszás, illetve a kevésbé likvid piacokon a nagyobb kereskedések esetében az árra gyakorolt hatás, jelentenek azonnali kihívást, amelyet egy ilyen stratégia sikere előtt meg kell oldani.

Az árak hatásának tanulmányozása a gépi tanulás teljesen más paradigmáját indokolja - az offline tanulás helyett.¹¹A modell egy RL-ügynök, amely képes frissíteni optimális politikáját, amint új információk érkeznek, és új, a jövőbeli információkat befolyásoló cselekvéseket hajtunk végre.

A következő példában azt a végrehajtási problémát vizsgáljuk, hogy hogyan kell a blokkos eladási megbízásokat felbontani a kockázattal korrigált hozam maximalizálása érdekében. A példában a részvényárfolyamot szimuláljuk, nem pedig megjósoljuk, és a legjobb megközelítést a részvényárfolyamtól függő ütemezési problémaként próbáljuk meghatározni. A probléma megfogalmazása ezért nagyobb kihívást jelent, és nagyobb szakértelmet igényel a készen kapható gépi tanulás alkalmazása ilyen problémákra.

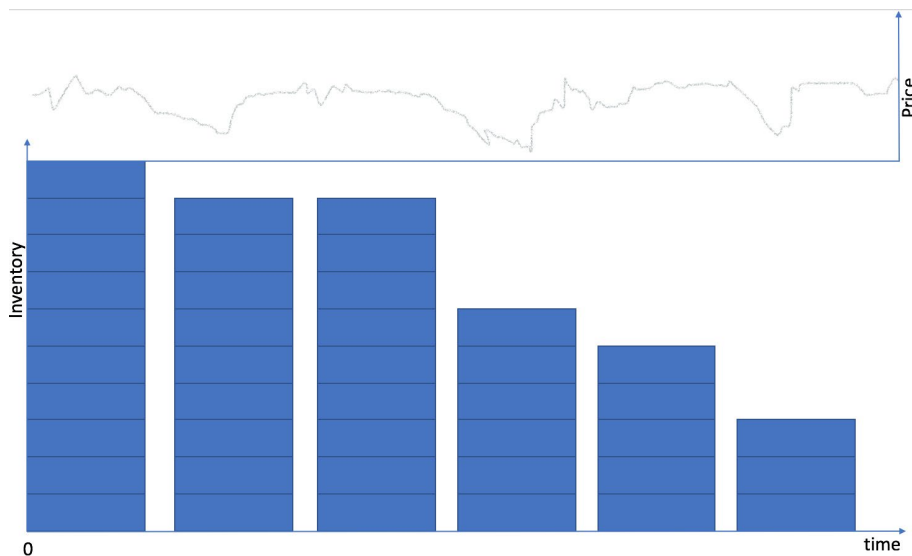
5.3 Optimális részvénykivitelezés megerősített tanulással

Tegyük fel, hogy egy brókernek N részvénycsomagot kell eladnia, amelyekben minden blokkban n részvény van, például $N = 10$, $n = 1000$. A probléma az, hogy hogyan osszuk fel az N blokkot, hogy maximalizáljunk valamilyen jutalmat, amint azt az ábra szemlélteti. 9.

A megerősítéssel tanulás alkalmazása során a hangsúlyt a következők meghatározására helyezzük: (i) az állapot - gyakran egy megfigyelt változó; (ii) a cselekvések, amelyek az ágens által hozott állapotfüggő döntések; és (iii) a döntésekért járó jutalmak.

A készlet t időpontban fennálló állapotát az \mathbf{X}_t változó adja meg, amely egy $N = 10 \times n$ állapotú halmazban \mathbf{X} értékeket vesz fel, így a $t = 0$ időpontban a kiindulási pont $0\mathbf{X}_0 = \mathbf{X}^{(N-1)}$, a célállapot pedig $\mathbf{X}_T = \mathbf{X}^{(0)} = 0$, azaz a bróker a teljes készlet eladására törekszik. Minden egyes lépésben az ügynöknek négy lehetséges $\mathbf{a}_t = \mathbf{a}$ akciója van, ⁽ⁱ⁾amelyek a t időpontban eladott részvényblokkok számát mérik, ahol $\mathbf{a}^{(0)} = 0$ nem akció, és $\mathbf{a}^{(i)} = i$, $i = 1, \dots, 3$ eladott tömbök száma. A

¹¹Ez alól kivételt képeznek a Bayes-féle felügyelt tanulási módszerek, például a változáspont-felismerés, amelyek képesek frissíteni a poszterior értéket az új adatokra reagálva.



9. ábra: Az optimális végrehajtás problémája: hogyan lehet a nagy piaci megbízásokat kisebb, kisebb piaci hatású megbízásokra bontani. Az állapotérnek a leltárt választjuk, amelyet a blokkok száma, a részvényárfolyam és az idő mutat. Ezen az ábrán az ügynök minden egyes időlépésben eldönti, hogy 0, 1, 2, ..., blokkokat ad-e el. A probléma tehát az, hogy megtartsa-e a készletet, növelve a piaci kockázatot, de csökkentve a piaci hatást, vagy gyorsan eladja a készletet, hogy csökkentse a kockázatot, de növelje a piaci hatást.

frissítési egyenlet
a következő

$$X_{t+1} = (X_t - a_t)_+ \quad (2)$$

Feltételezzük, hogy az ügyletek lineáris piaci hatáson keresztül befolyásolják a részvényárfolyam dinamikáját.

$$s_{t+1} = Ste^{(1-vat)} + \sigma StZ_t \quad (3)$$

ahol v egy piaci súrlódási paraméter. Egy véges Markov-döntési folyamat problémájára való leképezéshez a lehetséges részvényárak tartománya M értékre diszkrétizálható, pl. $M = E12$. lehetővé teszi az optimális politika táblázatos formában történő megjelenítését - ami kényelmes a kimenet értelmezéséhez.¹² A probléma állapotterének dimenziója $N \times M = 10 \cdot 12 = 120$. A kiterjesztett tér dimenziója az idővel együtt tehát $120 \cdot 10 = 1200$.

Egy t részvénycsomag eladásának kifizetése, ha a részvények árfolyama S_t , naS_t . A kockázattal korrigált kifizetés a következő $t + 1$ lépésnél a fennmaradó készletár varianciájára vonatkozó büntetést ad hozzá: $r_t = naS_t - \lambda nVar [SX_{t+1t+1}]$. Az állapot és az idő minden kombinációja egy háromdimenziós rácsként ábrázolható, amelynek mérete $N \times M \times T = 10 \cdot 12 \cdot 10$. Minden egyes időlépésnél tehát egy időfüggő optimális politikát találunk 10×12 (a készlet és a készletárszint esetében) állapotokkal és négy lehetséges $a_t = a, a_0, a_{12}, a_3$ akcióval, amely egy $10 \cdot 12$ mátrixnak tekinthető, amint az a második időlépésre a táblázatban látható. 1.

¹²Továbbá, ha az állapotter diszkrét, akkor olyan egyszerű RL módszereket használhatunk, mint a SARSA és a Q-tanulás, szemben a bonyolultabb módszerekkel, mint a policy gradiens vagy a folyamatos állapotú Q-tanulás, függvény közelítéssel (pl. mély Q-tanulás).

Most már alkalmazhatjuk a SARSA-t az árhatás-probléma megoldásának megtanulására egy ilyen egyszerűsített környezetben. Az on-line tanuláshoz szükséges feltárára ϵ -greedy politikát használhatunk. Az ϵ -greedy politika egy egyszerű sztochasztikus politika, amelyben az ágens olyan akciót hajt végre, amely ϵ valószínűséggel maximalizálja az akció-érték függvényt, és ϵ valószínűséggel egy tisztán véletlenszerű akciót hajt végre.

Az ϵ -greedy politikát mindkét a, a' akció előállítására használjuk a SARSA frissítésben, míg a Q-tanulás esetén csak az aktuális lépés akciójának kiválasztására. Megfelelően kis α esetén és az ϵ kúposítása mellett (lásd a 10. ábrát) mindkét módszer az ábrán láthatóan ugyanahhoz a kumulatív jutalomhoz 11konvergál.

Az optimális politika a $t =$ időpontban a SARSA segítségével a kereskedés végrehajtási problémára. A sorok a készletszintet, az oszlopok pedig a készletárszintet jelölik. Minden elem egy $\{0, 1, 2, 3\}$ részvénycsomag eladására irányuló műveletet jelöl. Megjegyezzük, hogy minden egyes időlépéshez más-más táblázatos mátrix áll rendelkezésre, de a bemutatás rövidsége miatt itt nem mutatjuk be.

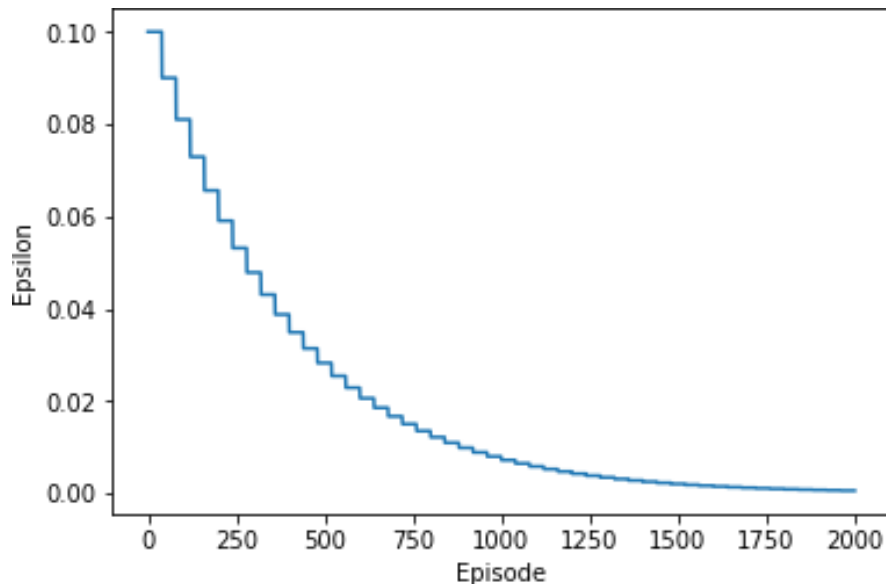
$t = 2$	Árszint											
Leltár	1	2	3	4	5	6	7	8	9	10	11	12
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0	1	0	0	0	0	0	0	0	0	0	0
5	0	3	1	0	0	0	0	0	0	0	0	0
6	0	3	2	3	0	0	0	0	0	0	0	0
7	0	3	2	2	0	0	0	0	0	0	0	0
8	0	0	0	2	1	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0

1. táblázat: Az optimális politika a $t = 2$ időbeli lépésben a SARSA használatával a kereskedés végrehajtási problémára. A sorok a készletszintet, az oszlopok pedig a készletárszintet jelölik. Minden elem a $\{0, 1, 2, 3\}$ részvényblokkok eladására irányuló műveletet jelöli.

Bár a probléma megfogalmazása egyszerűbb, mint amire a termelési szintű alkalmazásokhoz szükség van, a példánk arra szolgál, hogy szemléltesse az értelmezhetőség fontosságát és a különböző módszerek - ebben az esetben a Q-learning és a SARSA - összehasonlításának lehetőségét, amelyek köztudottan konvergálnak, amint ϵ nullához közelít. Általában a Q-learninget részesítjük előnyben, mivel ez a módszer politikán kívüli, és nem feltételezi, hogy az optimális politika generálta az adatokat.

A fenti példa a diszkrét állapotú és diszkrét cselekvéses megerősítéses tanulás egyszerű illusztrációját nyújtja. Sok probléma esetén a folytonos állapotokban és/vagy cselekvésekben való gondolkodás előnyösebb lehet. Ilyen esetekben függvényközelítéseket kell használni a Q-függvény vagy/és a cselekvési politika reprezentálására. Amint azt Halperin (2017); Halperin (2019) bemutatja, a folytonos állapotú Q-tanulás felhasználható az opciós árazás modellfüggetlen módjának biztosítására, amely a modellkockázat mérséklésén túlmenően a klasszikus Black-Scholes-modell (és más "kockázatsemleges" modellek) javításával a diszkrét időben történő téves fedezeti kockázat kezelésével is javít a klasszikus Black-Scholes-modellen (és más "kockázatsemleges" modelleken).¹³ Ha egy egyszerű kvadratikus Markowitz-típusú hasznossági függvényt használunk ebben a beállításban, a Q-függvény lineáris függvény közelítése elegendő ahhoz, hogy

¹³Ez a kockázat eltűnik az eredeti BS modellből és az összes "kockázatsemleges" modellből, ami annak köszönhető, hogy ezek a modellek konstrukciójuknál fogva folytonos időben vannak megfogalmazva, ahelyett, hogy a folytonos idejű határértéket a diszkrét idejű dinamika közelítéseként kapnánk.



10. ábra: Ez az ábra azt szemlélteti, hogy az ϵ hogyan csökken az árra gyakorolt hatás problémájában növekvő epizódokkal úgy, hogy a Q-tanulás és a SARSA ugyanahhoz az optimális politikához és kumulatív jutalomhoz konvergál.

mind az optimális opciós árat, mind az optimális fedezetet megkapja. Bonyolultabb hasznossági függvények, pl. CVaR hasznosság esetén bonyolultabb függvényközelítéseket, például neurális hálózatokat használhatunk. Ez mély megerősítéses tanulást jelent (pl. mély Q-tanulás), amely kombinálja a felügyelt tanulást és a megerősítéses tanulást. A Q-tanulás G-tanulásnak nevezett, entrópiával szabályozott változatát használó, piaci hatáshatással járó portfóliókonstrukciókra való kiterjesztéseket a (Halperin and Feldshteyn, 2018) írja le. Az RL más klasszikus pénzügyi problémákra, például a vagyonkezelésre és a pénzügyi tervezésre való alkalmazásáról lásd a hamarosan megjelenő könyvet (Bilokon et al., 2020).

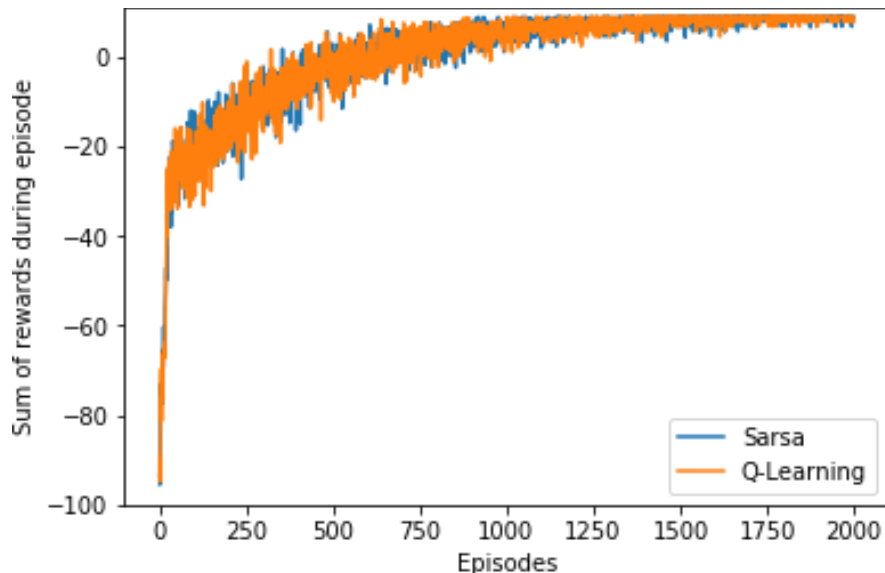
További érdekes irányokat kínál az inverz megerősítéses tanulás (IRL), amelynek célja, hogy megtalálja egy ágens jutalomfüggvényét a megfigyelt viselkedése alapján. A konkrét kereskedők jutalmazási (hasznossági) függvényeinek tanulására vonatkozó alkalmazások mellett az IRL alkalmazható a piac "láthatatlan kezét" megtestesítő, egész piacra kiterjedő "ágensre" is, azaz az IRL használható a piac egészének dinamikájának modellezésére (Halperin és Feldshteyn, 2018).

Érdekes módon a piaci dinamika így kapott modellje implicit módon feltételezi, hogy a piac jóindulatú rezsimben van, ahol nincsenek vállalati csődök vagy válságok. A statisztikai és kvantumfizikából származó módszerek alkalmazásával Halperin és Dixon (2018) az így kapott

modellt kiterjesztette a piaci összeomlásokra és vállalati csődökre is - olyan jelenségekre, amelyek mindkettő nagyon fontos, de a részvénytörzsek standard modelljeiből hiányoznak (hacsak nem vezetünk be speciális szabadságfokokat, például sztochasztikus hitelfelárakat).¹⁴

Megjegyezzük, hogy ez utóbbi munka bizonyos értelemben "túlmutat az IRL-en" (vagy általában az ML-en, ha már itt tartunk), mert megmutatja a ritka adatok (vagy akár a megfigyeletlen adatok!) fontosságát egy megfelelő dinamikus környezeti modell felépítésében. Míg a tiszta ML algoritmusokat az adatok, a korlátok és a korlátozások vezérik.

¹⁴E témák bevezető ismertetése, valamint egy általános bevezetés a megerősítéses tanulásról, az opcióárazási és portfólióoptimalizálási példákkal együtt, megtalálható egy Coursera specializációban (Halperin, Halperin).



11. ábra: A Q -tanulás és a SARSA a megfigyelések szerint ugyanahhoz az optimális politikához és kumulatív jutalomhoz konvergál az árra gyakorolt hatás problémájában az ϵ -csapolásos ütemezés szerint, ahogy az ábrán látható. 10. Megjegyzendő, hogy a kumulatív jutalmak húsz szimuláció átlagát jelentik.

vagy a világ "igaz" természetére vonatkozó hiedelmek megfelelő regularizációs büntetések alkalmazásával érvényesíthetők - amelyek maguk is alapulhatnak modelleken, lásd Halperin and Dixon (2018) egy példáját. Összefoglalva, a szekvenciális döntéshozatal problémáihoz, amelyekben az ágens cselekvéseiből lehetséges visszacsatolási hurokkal rendelkeznek, a megerősítő tanulási algoritmusok gyakran a legtermészetesebb és legátfogóbb megközelítést nyújtják. Az állapottér, az akciók, a jutalmak és a felfedezés kúpjának megválasztásának képessége a legfontosabb szempontok közé tartozik, hogy a projektnek megfelelő ütemezést kell követnie, és védhető eredményeket kell produkálnia.

Általános ajánlásunk az, hogy kerüljük a mély megerősítő tanulás elhamarkodott alkalmazását a "sekély" megerősítő tanulás első alkalmazása előtt - legalább egy viszonyítási alap létrehozása a fejlettebb technikákhoz, például a mély Q -tanuláshoz.

6 Összefoglaló

A gépi tanulás általánosabb keretet biztosít a pénzügyi modellezéshez, mint lineáris paraméteres elődei, általánosítva az olyan archetipikus modellezési megközelítéseket, mint a faktormodellezés, a származtatott termékek árazása, a portfólióépítés és -fogyasztás, az optimális fedezés modellmentes, adatvezérelt megközelítésekkel, amelyek robusztusabbak a modellkockázattal szemben és megragadják a kiugró értékeket.

Ebben a rövid állásfoglalásban értékeltük az elfogadás néhány felmerült akadályát. - a legtöbbjük a szociológia eme interdiszciplináris területének műtárgya. Azonosítjuk a főbb vészjelzőket, és iránymutatásokat és megoldásokat fogalmazunk meg ezek elkerülésére. A felügyelt tanulást és a megerősítést a befektetési menedzsmentben és a kereskedésben alkalmazó példákat mutatjuk be a gépi tanulás sikeres alkalmazásának legjobb gyakorlataihoz.

Az NLP és a szövegbányászat alkalmazásain túl a gépi tanulásnak központibb szerepe van a pénzügyi modellezésben. A sikeres alkalmazás kulcsa az új és a régi ötvözése, a gépi tanulás pénzügyi modellezési és ökonometriai keretrendszerekbe való beillesztése, amelyek általánosítják klasszikus társaikat, és így átveszik az értelmezhetőséget és a robusztusságot, ami szükséges ahhoz, hogy a gépi tanulás a pénzügyekben sikeres legyen.

Hivatkozások

- Alonso, M. N., G. Batres-Estrada és A. Moulin (2018). *Mélytanulás a pénzügyekben: Prediction of Stock Returns with Long Short-Term Memory Networks*, 13. fejezet, pp. 251-277. John Wiley Sons, Ltd.
- Arnold, V. I. (1957). Három változó függvényeiről. Kötet pp114., 679-681.
- Bayer, C. és B. Stemper (2018, október). Deep calibration of rough stochastic volatility models. *arXiv e-prints*, arXiv:1810.03399.
- Bilokon, P., M. Dixon és I. Halperin (2020). *Gépi tanulás a pénzügyekben: az elmélettől a gyakorlatig* (1. kiadás). Springer.
- Borovykh, A., S. Bohte, and C. W. Oosterlee (2017, Mar). Conditional Time Series Forecasting with Convolutional Neural Networks. *arXiv e-prints*, arXiv:1703.04691.
- H. H., L. Gonon, J. Teichmann, and B. Wood (2018, Feb). Deep Hedging. *arXiv e-prints*, arXiv:1802.03042.
- Carvalho, C. M., H. Lopes, O. Aguilar és M. Mendoza (2012, 01). Dinamikus részvényválasztási stratégiák: Egy strukturált faktormodell keretrendszer. *Bayesian Statistics 9*.
- Chen, L., M. Pelger és J. Zhuz (2019, március). Mélytanulás az eszközárzásban. Technikai jelentés, Stanford Egyetem.
- de Prado, M. (2018). *Advances in Financial Machine Learning*. Wiley.
- Dhar, V. (2013, december). Adattudomány és előrejelzés. *Commun. ACM (5612)*, 64-73.
- Dixon, M. (2018). Egy nagyfrekvenciás kereskedés végrehajtási modellje felügyelt tanuláshoz. *High Frequency (11)*, 32-52.
- Dixon, M., D. Klabjan és J. H. Bang (2016). Osztályozáson alapuló pénzügyi piaci előrejelzés mély neurális hálózatok segítségével. *CoRR abs/1603.08604*.
- Dixon, M. F. és N. G. Polson (2019, március). Deep Fundamental Factor Models. *arXiv e-prints*, arXiv:1903.07677.
- Fama, E. és J. D. MacBeth (1973). Kockázat, hozam és egyensúly: Empirikus tesztek. *Journal of Political Economy (813)*, 607-36.
- Fama, E. F. és K. R. French (1992). A várható részvényhozamok keresztmetszete. *The Journal of Finance (472)*, 427-465.

- Fama, E. F. és K. R. French (1993). Közös kockázati tényezők a részvények és kötvények hozamában. *Journal of Financial Economics* (331), 3- 56.
- Fama, E. F. és K. R. French (2015). Ötfaktoros eszközárzási modell. *Journal of Financial Economics* (1161), 1-22.
- Feng, G., J. He, and N. G. Polson (2018, Apr). Deep Learning for Predicting Asset Returns. *arXiv e-prints*, arXiv:1804.09314.
- Flood, M., H. V. Jagadeh, and L. Raschid (2016). A nagyméretű adatokkal kapcsolatos kihívások és lehetőségek a pénzügyi stabilitás nyomon követésében. *Financial Stability Review* (20), 129-142.
- Gu, S., B. T. Kelly és D. Xiu (2018). Empirikus eszközárzások gépi tanúlással. Chicago Booth Research Paper 18-04.
- Guida, T. (2019). *Big Data és gépi tanulás a kvantitatív befektetésekből*. Wiley Finance. Wiley.
- Halperin, I. Gépi tanulás és megerősített tanulás a pénzügyekben. *Coursera specializáció*, <https://www.coursera.org/specializations/machine-learning-reinforcement-finance>.
- Halperin, I. (2017, december). QLBS: Q-Learner in the Black-Scholes(-Merton) Worlds. *megjelenés alatt: Journal of Derivatives*, arXiv:1712.04609.
- Halperin, I. (2019). A QLBS Q-Learner goes NuQLear: illesztett Q iteráció, inverz RL és opciós portfóliók. *Quantitative Finance* (199), 1543-1553.
- Halperin, I. és M. F. Dixon (2018). "Quantum Equilibrium-Disequilibrium": Asset Price Dynamics, Symmetry Breaking, and Defaults as Dissipative Instantons. *forthcoming in Physica A abs/808.03607*, arXiv:1808.03607.
- Halperin, I. és I. Feldshteyn (2018, május). A jelzések piaci öntanulása, hatása és optimális kereskedés: Invisible Hand Inference with Free Energy (Or, How We Learned to Stop Worrying and Love Bounded Rationality). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3174498.
- Harvey, C. R., Y. Liu, and H. Zhu (2015, 10). . . és a várható hozamok keresztmetszete. *The Review of Financial Studies* (291), 5-68.
- Heaton, J. B., N. G. Polson és J. H. Witte. Mélytanulás a pénzügyekben: mély portfóliók. *Applied Stochastic Models in Business and Industry* (331), 3-12.
- Kolmogorov, A. N. (1957). A sok változóból álló folytonos függvények ábrázolásáról egy változóból álló folytonos függvények szuperpozíciójával és összeadással. *Dokl. Akad. Nauk SSSR* 114, 953-956.
- Martin, C. H. és M. W. Mahoney (2018). Implicit önregularizáció mély neurális hálózatokban: Evidence from random matrix theory and implications for learning. *CoRR abs/1810.01075*.
- Martin, D. (1977). A bankcsőd korai előrejelzése: A logit regressziós megközelítés. *Journal of Banking & Finance* (13), 249-276.

- Moritz, B. és T. Zimmermann (2016). Fa alapú feltételes portfólióválogatás: A múltbeli és jövőbeli részvényhozamok közötti kapcsolat.
- Rosenberg, B. és V. Marathe (1976). Közös tényezők az értékpapírhozamokban: Mikrogazdasági de-terminanciák és makrogazdasági korrelációk. Research Program in Finance Working Papers 44, University of California at Berkeley.
- Sirignano, J., A. Sadhwani, and K. Giesecke (2016, július). Deep Learning for Mortgage Risk (Mélytanulás a jelzalogkockázathoz). *ArXiv e-prints*.
- Swanson, N. R. és H. White (1995). Modellválasztásos megközelítés a terminusszerkezet információinak értékelésére lineáris modellek és mesterséges neurális hálózatok segítségével. *Journal of Business & Economic Statistics* (133), 265-275.
- Trippi, R. és D. Desieno (1992, 01). Részvényindex határidős ügyletek kereskedése neurális hálózattal. *19*, 27-33.
- van den Oord, A., S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbren-ner, A. W. Senior, and K. Kavukcuoglu (2016). Wavenet: Egy generatív modell a nyers hangok számára. *CoRR abs/1609.03499*.

A statisztikai vizsgálatok összefoglalása

Név	Leírás
Augmentált Dickey-Fuller teszt	Annak megállapítására szolgál, hogy az idősoros adatok stacionáriusak.
Ljung-box teszt	Annak meghatározására szolgál, hogy az adatok korreláltak vagy i.i.d.
Granger-okozati teszt	A bemeneti változók azonosítására szolgál (más néven jellemzők) a felügyelt tanuláshoz.

2. táblázat: Összefoglaló a pénzügyek felügyelt gépi tanulásának néhány leghasznosabb feltáró adatelemzési tesztjéről.

Név	Leírás
Chi-négyzet teszt	Annak meghatározására szolgál, hogy egy osztályozó zavarmátrixa statisztikailag szignifikáns, vagy csupán fehér zaj.
t-test	Annak meghatározására szolgál, hogy két különálló regressziós modell kimenete i.i.d. adatokon statisztikailag különböznek.
Mariano-Diebold teszt	Annak meghatározására szolgál, hogy két különálló idősoros modellek statisztikailag különböznek.
Fehér teszt	Annak meghatározására szolgál, hogy a hiba heteroszkedasztikus vagy homoszkedasztikus
Portmanteau teszt	Általános teszt arra vonatkozóan, hogy a hiba egy idősoros modellben autokorrelált. A Box-Ljung és a Box-Pierce tesztek példái.

3. táblázat: *Összefoglaló a pénzügyekben alkalmazott felügyelt gépi tanulás néhány leghasznosabb diagnosztikai tesztjéről. A felügyelt gépi tanulás példája az idősoros módszereknél a rekurrens neurális hálózatok.*