[Feladat 2](#_Toc497333920)

[1 Magyar áramtőzsde 4](#_Toc497333921)

[2 Tradicionális statisztikai módszerek 11](#_Toc497333922)

[2.1 Idősorok 11](#_Toc497333923)

[2.2 Trend elemzési modellek 11](#_Toc497333924)

[2.3 Regresszió 12](#_Toc497333925)

[2.3.1 Lineáris regresszió 12](#_Toc497333926)

[Kétváltozós eset 12](#_Toc497333927)

[2.3.1.1 Becslési módszerek pl.: 12](#_Toc497333928)

[2.3.1.2 Legkisebb négyzetek módszere 12](#_Toc497333929)

[2.3.1.3 Elaszticitás 13](#_Toc497333930)

[2.4 Mozgó átlagolás 13](#_Toc497333931)

[2.4.1 Egyszerű mozgóátlag 14](#_Toc497333932)

[2.4.2 Exponenciális simítás, 14](#_Toc497333933)

[2.4.2.1 Egyszeres exponenciális simítás 14](#_Toc497333934)

[2.4.2.2 Brown féle kettős exponenciális simítás 15](#_Toc497333935)

[2.5 Legközelebbi szomszéd 15](#_Toc497333936)

[2.6 ARMA modell 16](#_Toc497333937)

[2.7 Modellek összehasonlítása: 17](#_Toc497333938)

[3 Neurális hálók 19](#_Toc497333939)

[4 Irodalom jegyzék: 21](#_Toc497333940)

Feladat

Idősor-előrejelzésre alkalmas módszerek áttekintése, implementálása, összehasonlítása, használhatóságuk feltérképezése. A módszerek igen széles spektrumát használták, javasolták már erre a célra a legközelebbi szomszéd alapú modellek, ARMA, ARIMA modellek, mesterséges neurális hálók, regressziós modellek stb. ezek áttekintése és egy csoport kiválasztása, majd megvalósítása és tesztelése a feladat. Részben szimulált, részben valós, mért adatokon célszerű tesztelni a megvalósított módszereket.

A magyar áramtőzsde (HUPX) 2009-es indulása óta már több évnyi klíringeredmény áll rendelkezésre mélyrehatóbb elemző vizsgálatok elvégzéséhez.

**Klíringrendszer**: kölcsönös leszámítási rendszer, amelyben az elszámolási viszonyban levő pénzintézetek tartozásaikat kiegyenlítik. Készpénz nélküli elszámolási rendszer, amikor készpénzben vagy valutában csak az egyenleget, vagyis a tartozások és követelések különbözetét, a klírincsúcsot egyenlítik ki. Célja a pénzforgalom csökkentése (leszámolás). Belföldi viszonylatban, a bankügyletekben használatos. A külkereskedelemben a részt vevő kereskedők csak saját bankjukkal állnak kapcsolatban, velük hazai törvényes fizetési eszközben számolnak el. Nemzetközi szinten általában két ország jegybankja között szokásos (bilaterális klíring) vagy több ország közötti elszámolás központi rendezése (multilaterális klíring). A klíringcsúcs a klírinelszámolásokban az év végén keletkező egyenleg, amelynek kiegyenlítése történhet pótlólagos áruszállítással, arannyal vagy konvertibilis valutával, illetve jegybanki hitellel.

Az elmúlt 8 évben ráadásul háromféle környezetben is működött a magyar tőzsde: önállóan, a cseh-szlovák-magyar összekapcsolásban, valamint a 4M piac-összekapcsolásban. 21[1]

**4M piac:** 4M piac-összekapcsolás – 4M Market Coupling (4M MC) Az Európai Parlament és a Tanács 2009. július 13-i 714/2009/EK rendelete a villamos energia határokon keresztül történő kereskedelme esetén alkalmazandó hálózati hozzáférési feltételekről, és az 1228/2003/EK rendelet hatályon kívül helyezéséről szóló rendelettel összhangban a regionális szinten harmonizált menetrend-kezelési folyamat és a koordinált kapacitásallokáció került bevezetésre. Ennek megfelően Magyarország 2012 szeptember 11től a Cseh Köztársasággal és Szlovákiával volt piaci-összekapcsolásban. Miután Románia is belépett az EU tagállamok közé így ugyancsak kötelezett lett a vele határos Közép-Kelet Európia régió országai számára előírt követelmények és elvárások teljesítésben. 2013 augusztusában a nemzeti energiaszabályozó hatóságok jóváhagyták a cseh-szlovák-magyar áram piac-összekapcsolást Romániával. Ennek következtében 4M piac-összekapcsolás 2014 november 19-i kereskedési nappal sikeresen elindult.

A 4M MC projekt technikai és gazdasági célja, hogy a funkciók központosításával, a bonyolult, sok szereplő közötti többoldalú viszonyokat egy- az - egy típusú kapcsolattá alakítsa, ezáltal elősegítve a piac-összekapcsolás további piacok irányába történő kibővítését. A megvalósítás előtt a felek megállapodtak egy olyan megoldás kialakításában, amely a lehető legnagyobb mértékben kompatibilis a nyugat-európai (Multi-Regional Coupling – MRC) régióval. Ezért a piac-összekapcsolás kiértékelési algoritmus bemeneti és kimeneti adatainak teljes kompatibilitását hozták létre az MRC-vel, lehetővé téve a két régió összekapcsolását, vagy a későbbiekben új felek csatlakozását. [2]

# Magyar áramtőzsde

A pénzügyi idősorok előrejelzését nagyban nehezíti, hogy ezek általában zajosak, nem stacionáriusak, nemlineárisak és kaotikusak, továbbá gyakran fordul elő bennük strukturális törés is Ezen okok miatt a pénzügyi/tőzsdei idősorok előrejelzése az egyik legnagyobb kihívás a piaci szereplők számára.

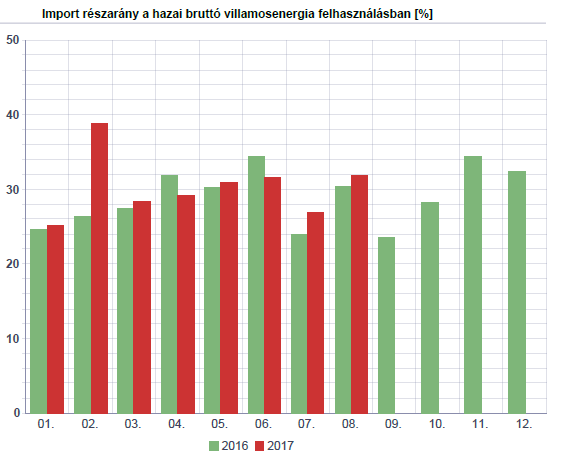
**MAVIR**: Magyar Villamosenergia-ipari Átviteli Rendszerirányító Zártkörűen Működő Részvénytársaság. A MAVIR – mint az átviteli hálózat tulajdonosa – felelős az átviteli hálózat folyamatos működtetéséért és fejlesztéséért. Összehangolja a magyar villamosenergia-rendszer és a szomszédos hálózatok működését (export-import). [3] [4]

A rendszerirányítás feladata az országos energiarendszer teljesítmény-egyensúlyának fenntartása, a mérlegkörök tervektől való eltéréseinek kiegyenlítése. Ehhez meghatározza a szükséges tartalékokat, a szabályozás számára lekötött teljesítményeket. Elkészíti a hálózatfejlesztési stratégiát és javaslatot tesz az erőműpark fejlesztésére.

A MAVIR jelentős szerepet játszik a piacszervezés területén, folyamatosan nyomon követi és alkalmazza a jogszabályi környezetet változását. A rendszerirányító alapította és működteti a magyar villamosenergia-tőzsdét (HUPX).

A MAVIR honlapról sok fajta adatot lehet beszerezni órás és 15 perces bontásban. Végül a terv és tény rendszerterhelést, az import-export adatokat, a határmetszéki áramlásokat és a széltermelést töltöttem le a teljes időszakra, de ezeken kívül is használtam a honlapról diagramokat, adatokat személtetés okán.

A MAVIR honlapról beszerzett dokumentumok és diagramok tanulmányozása során kiderült, hogy a hazai termelés nem elegendő a hazai kereslet kielégítésre. Az igényt átlagban 30% import révén tudják biztosítani. [5] Az alábbi diagramon (1. ábra) jól látszik az import aránya a hazai villamos energia felhasználásba.

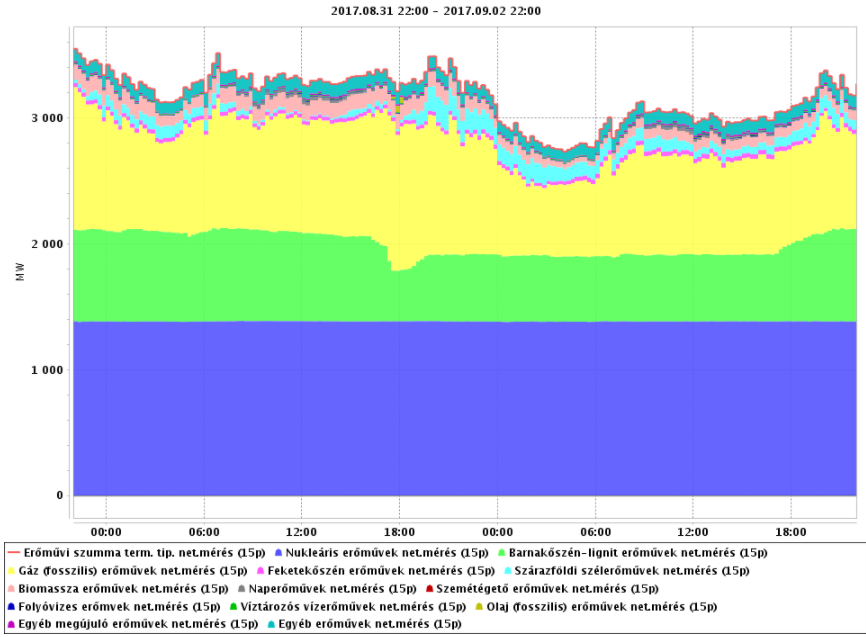


1. ábra

2016, 2017 évi import százalékos aránya a teljes rendszerterhelésben hónapos bontásban

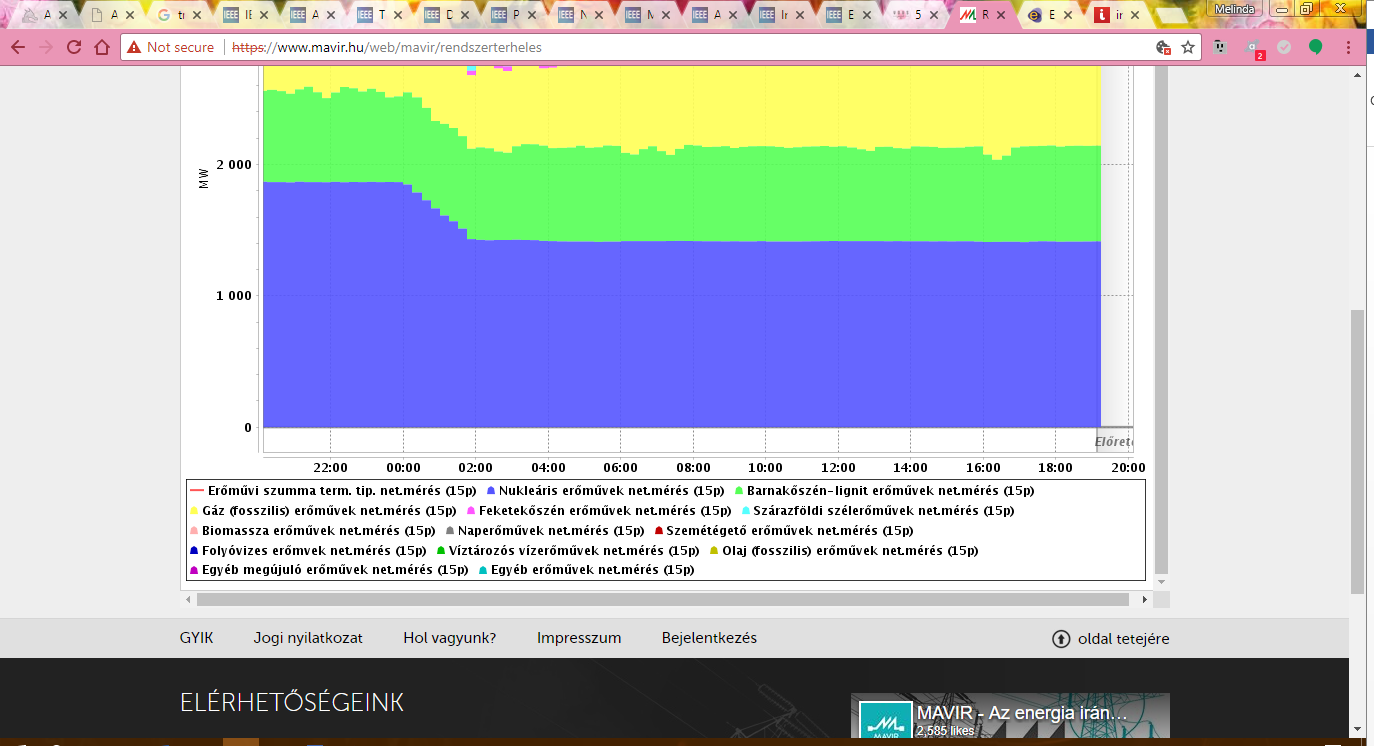
Érdekes hogy július és szeptember hónapban hírtelen csökken az import mennyisége. Illetve 2017 februárja nagyon kiugrott.

A maradék 70%-ot belső termelés adja. Többfajta erőmű szolgáltatja az áramot Magyarországon, az alábbi diagramon (2. ábra) 48 órás időszak látható. Itt jól látszik, hogy a paksi atomerőmű áramtermelése a legmeghatározóbb, ezek után a kőszén-, gázerőművek adják a legtöbb energiát és egy elenyésző rész a zöld energia, a víz, szél és egyéb megújuló energiák.

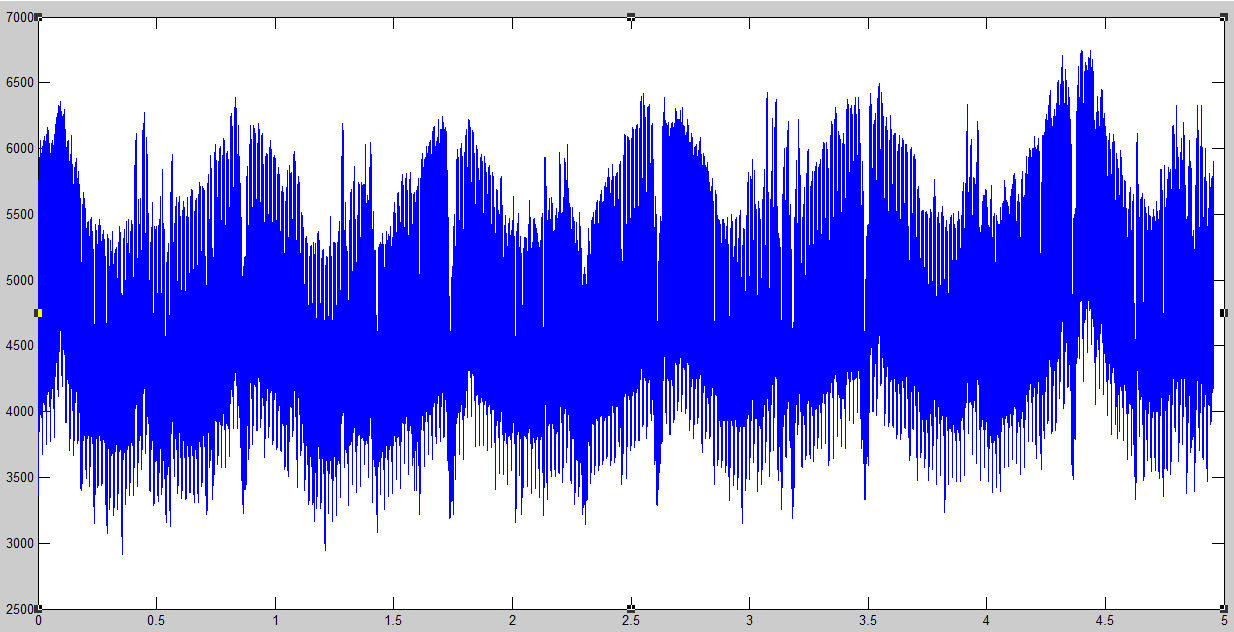


2. ábra

A magyarországi erőművek termelése, fajtánként órás bontásban, alább az egyes színek jelentése



2012 január 1 és 2017 szeptember 1 közötti adatokat gyűjtöttem be órás bontásban, ami 49680 mérést jelent. Meglepően jól látszik (3. ábra) az éves ciklus. Január, február, december több áramfogyasztás, nyáron leesik, bár az évközepén is vannak, kilengések gondolom ezek a hőség rekordos napok, illetve a decemberi hosszú ünnep karácsony és a szilveszter idején is jelentős változás figyelhető meg, de ebben az esetben visszaesés formájában, nem dolgozik senki. A 3. ábra a hitelesített tény rendszer terhelés mutatja, ennyi áramra volt igény az adott órában. Illetve az a trend is jól látszik, hogy évről évre több áramot fogyaszt a magyar nép.



3. ábra

Magyarországi rendszerterhelés órásbontásban 2012 január 1 és 2017 szeptember 1 között

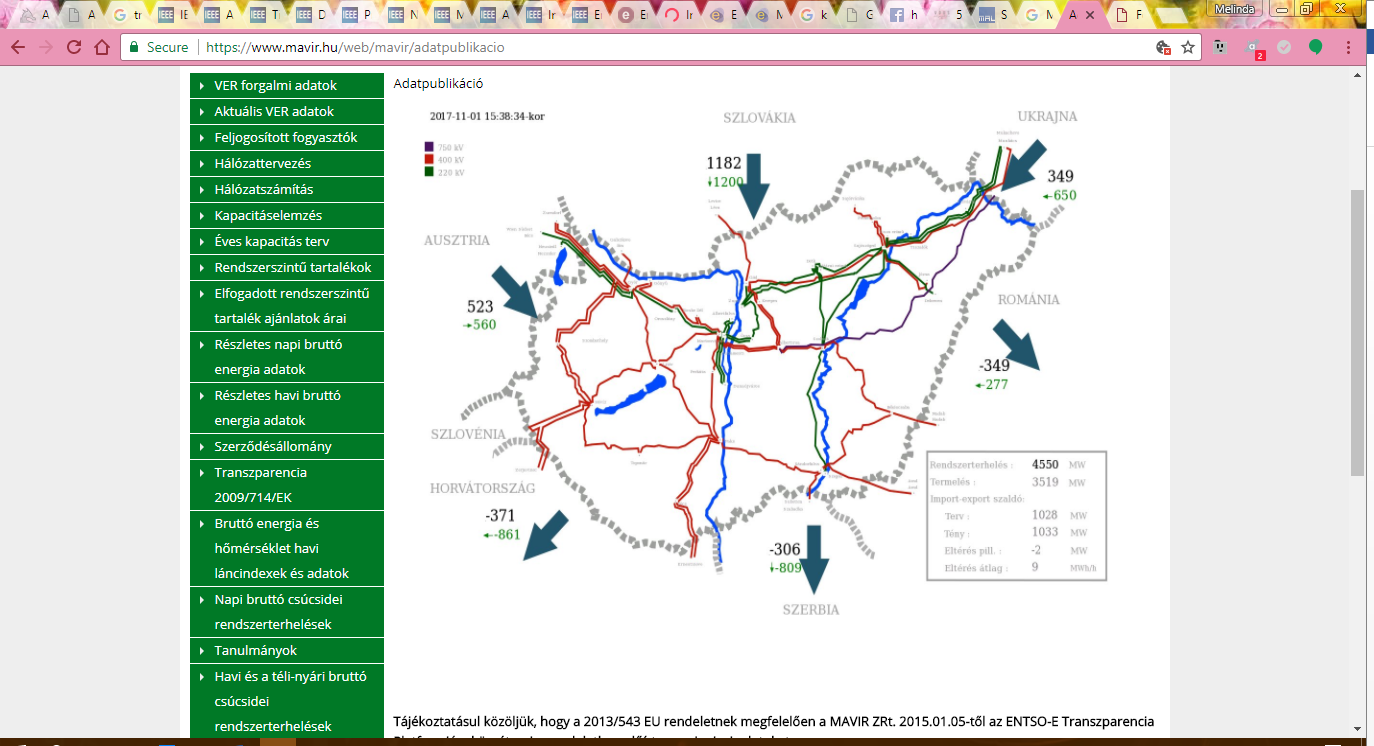
Sajnos nem teljesen hiánytalanok az adatok. Általában mindegyikből van egy terv és egy tényleges adat. A tényleges adatok sokszor üresek ilyenkor a tervezet adatokkal helyettesítem őket.

A másik nagy adatcsoport, amit le lehet tölteni a MAVIR-ról az a határmetszéki áramlások.

**Határmetszék**: a magyar és a szomszédos villamosenergia-rendszer közötti egyrendszerű vagy kétrendszerű távvezetékek, amelyek szinkron üzemmódban összekapcsolják a MAVIR és a szomszédos [rendszerirányító](http://www.enker.hu/glossary/term/285) szabályozási területeit. [6] A 4. ábra szemlélteti, hogy honnan mekkora áram folyik, folyhat be. Jól látszik, hogy:

* Magyarország – Ausztria HUAT
* Magyarország – Horvátország HUHR
* Magyarország – Románia HURO
* Magyarország – Szerbia HUSR
* Magyarország – Szlovákia HUSK
* Magyarország – Ukrajna HUUK

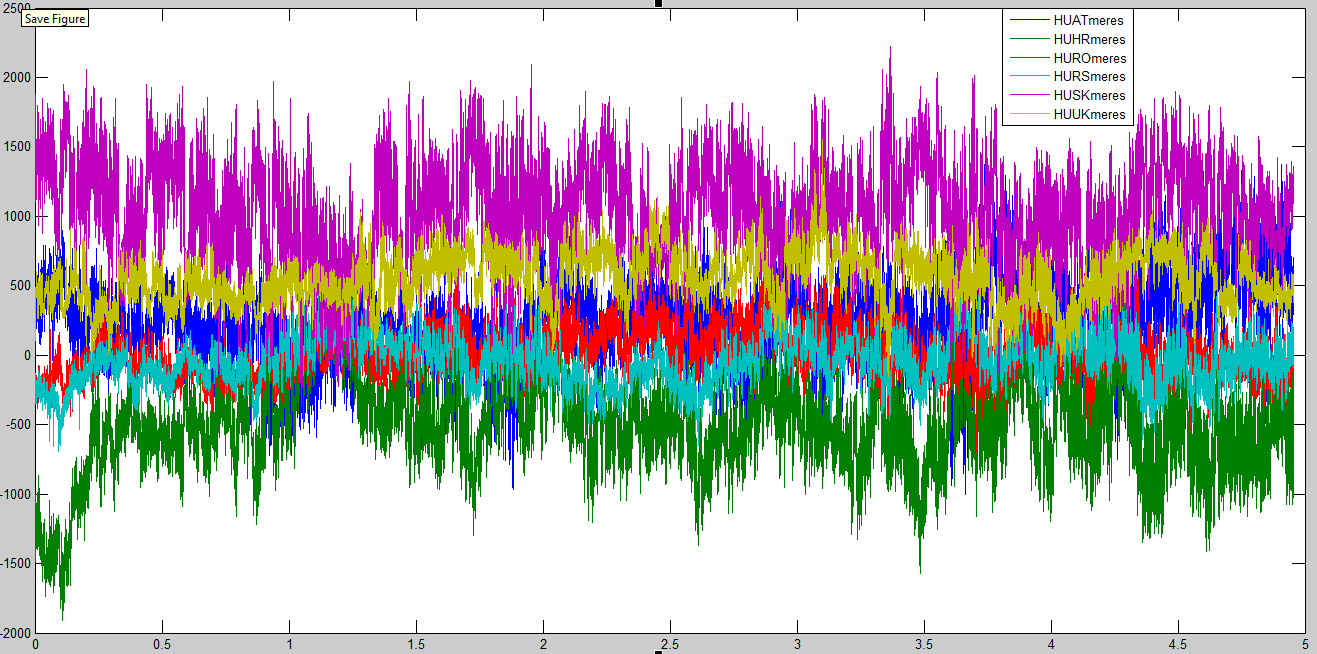
között van határmetszéki áramlás. Későbbiekben a 4 betűs rövidítéseiket fogom használni mikor rájuk hivatkozom. Jelentős határmetszék az HUUK mert egy 750kV egy 400 kV és 2 db 220 kV távvezetéken kapcsolódunk Ukrajnához, illetve jelentős még az ausztriai határmetszék és a szlovákiai, ezeken a pontokon több 400 kV-os illetve 220 kV-os távvezetéken kapcsolódunk. A következtetésemet alátámasztják még MAVIR honlapról beszerzett adatok, amelyek azt mutatják, hogy ezekből az országokból folyik be a legtöbb áram Magyarországra. A 4. ábra egy olyan Magyarország térkép, ami az aktuális pillanatban - jelen esetben 2017 november 1-e 15:38kor készült pillanatkép - mutatja a MAVIR honlapján a fennálló import-exportot és a rendszerterhelést. Éppen akkor, amikor készült a kép akkor Ausztria felöl 523 MW áram érkezett ténylegesen, míg a terv, ami zöld színnel van, 560 MW volt. Jól látszik, hogy az adott időpontban Szerbia felé áramot exportáltunk 306 MW, míg azt várta a MAVIR, hogy 809 MW fognak venni.



4. ábra

Határmetszéki áramlások pillanatnyi állapota 2017 November 1én 15:38:34kor

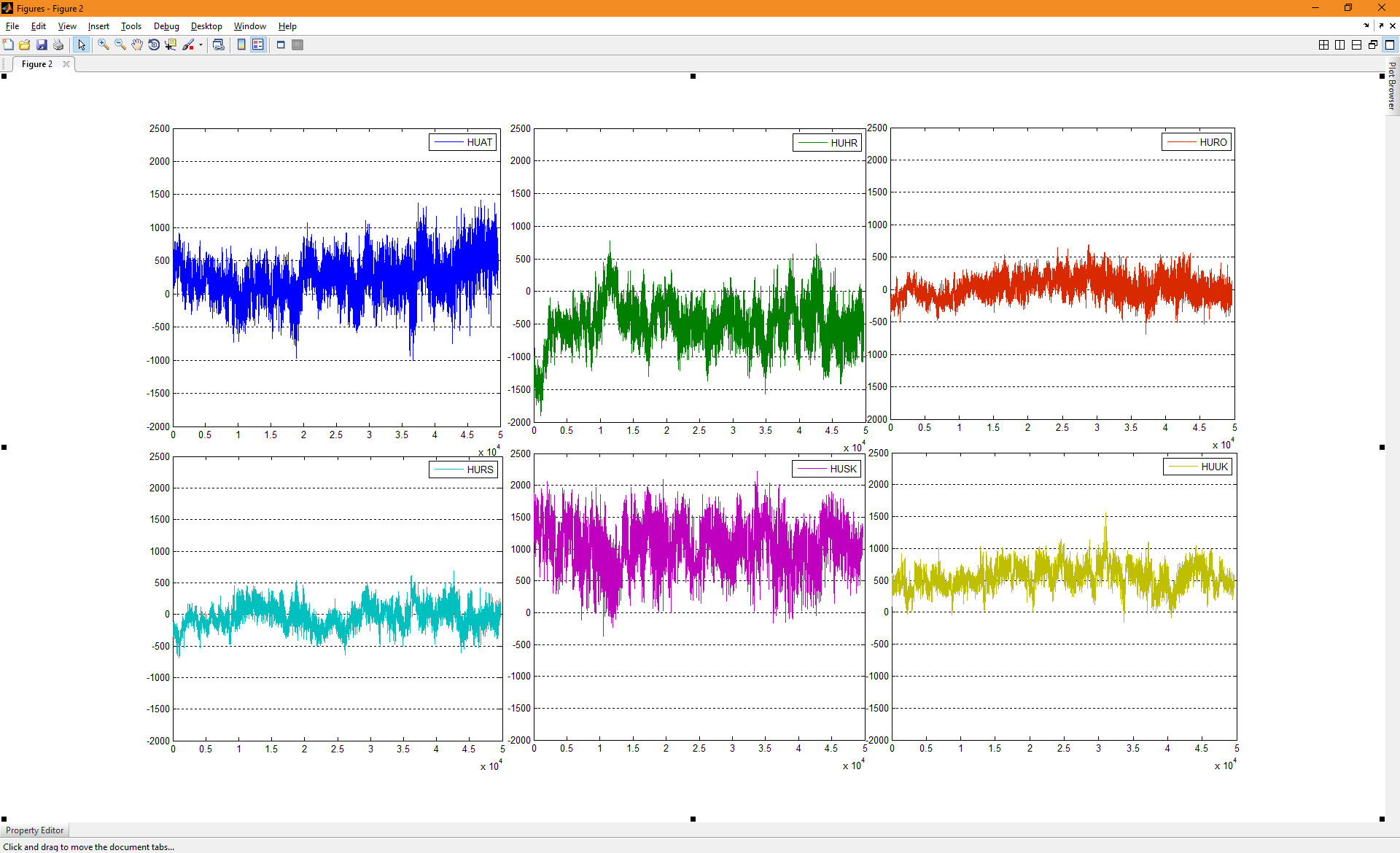
Nagyon fontos megjegyezni, hogy nem Magyarország használja fel az összes import áramot, Horvátország, Szerbia rajtunk keresztül részesednek a Közép-Európában termelt áramból. Az 5. ábra kaotikus színes világa mutatná be az egyes metszékekre jellemző exportot, importot.



5. ábra

Határmetszéki áramlások országos órás bontásban 2012 január 1 és 2017 szeptember 1 között

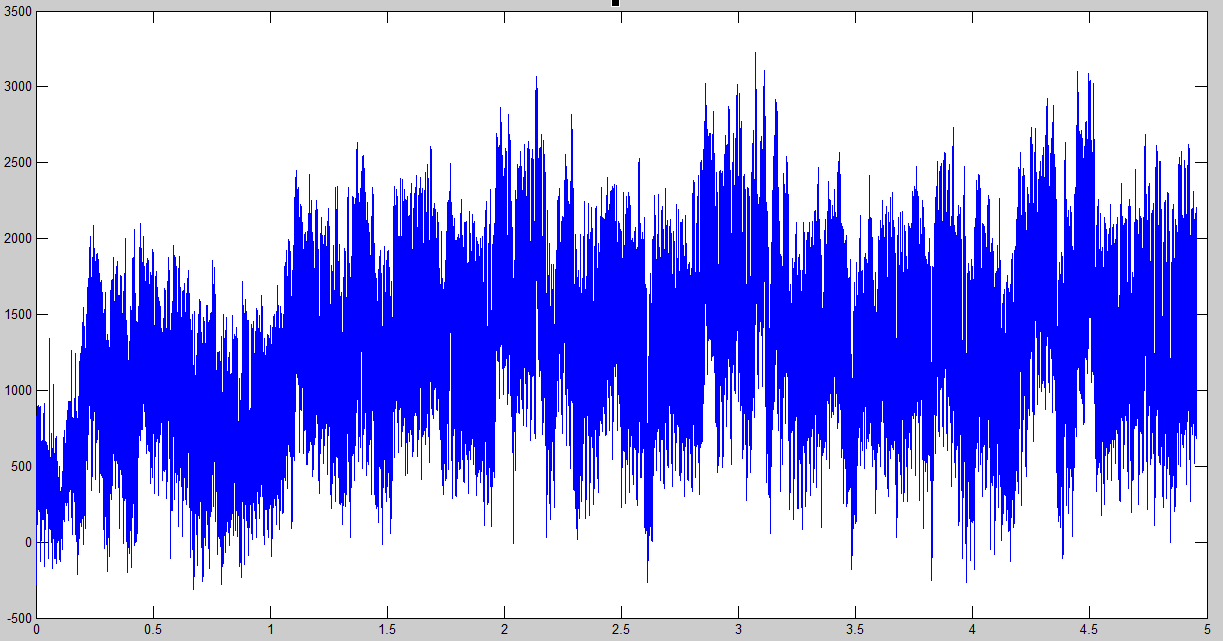
Az 5. ábra egymáshoz viszonyítva mutatja be az 6 db határmetszék export-importját. Már ezen is jó látszik, hogy a zöld színű HUHR többnyire negatív ami azt jelenti hogy Horváth országba főleg exportálunk áramot. A magenta színű HUSK egyértelműen pozitív az esetek döntőtöbbségében és kb. kétszerese a többi határmetszéknek, a letöltött adatok átlaga 1064,7 MW



6. ábra

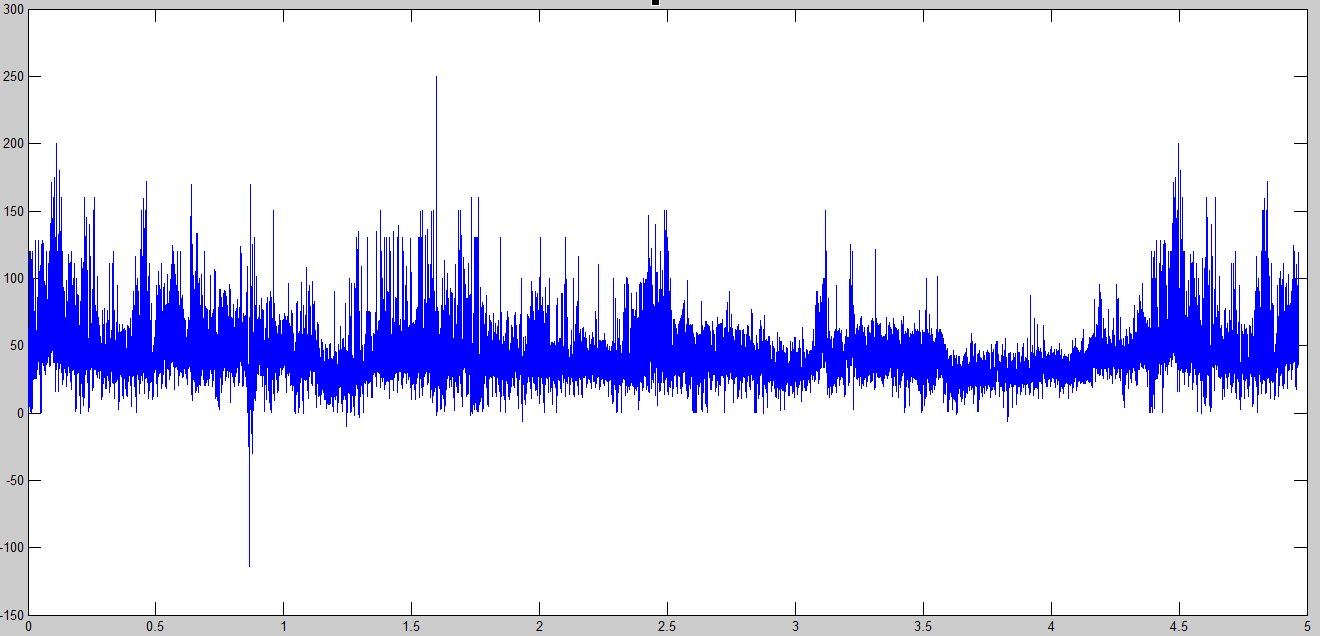
Határmetszéki áramlások országos bontása egymás mellett azonos skálával 2012 január 1 és 2017 szeptember 1 között

Ezek fényében megfogtam az összes határmetszéki adatot és summáztam őket, 7. ábra lett az eredmény, aminek maximuma 3224 kW lett, ami megdöbbentő annak a fényében, hogy a tény rendszerterhelés átlagosan 49330 kW. Nagyon szabálytalannak tűnik a rendszer terheléshez képest az export-import.



7. ábra

Export-import szaladó számított az országos adatok alapján



8. ábra

Árak órásbontásban 2012 január 1 és 2017 szeptember 1 között

# Tradicionális statisztikai módszerek

Abból indultam ki, hogy az áramtőzsde hasonlóan működik, mint a normál tőzsde, a kereslet és a kínálat határozza meg az árat, ezért tőzsdei algoritmusokat kerestem elsőnek, illetve megnéztem, hogy a matematikai tankönyvek, kiadott tananyagok milyen idősor elemzési elméleteket tartalmaznak.

A tözsdebarát [7] honlapon van egy lista a különböző elven működő, különböző módon használható előrejelzésekről. Leírás főleg azt tartalmazza, hogy ezen elvek használata mellett mikor érdemes adni-venni. De mivel én itt most nem fogok eladni, így pontosan, úgy ahogy le vannak írva nem használhatók számomra, ezért megkerestem az elvek pontos matematikai háttérét, erre főkképpen a [8] a számítástudományi és információelméleti tanszék által készített nagyméretű adathalmazok kezelése szeminárium anyagait használtam.

## Idősorok

Idősor megfigyelések egy sorozata, tipikusan időközönkénti mérés pl naponta, óránként … Nálam óránkénti mérés lesz. Számít a sorrend, időbélyeg tartozik az egyes elemekhez. Ki tudjuk használni, hogy az egymást követő megfigyelések erősen korrelálnak egymással. Az idősorokat megkülönböztetjük a változók száma szerint, lehet egyváltozós pl a hőmérséklet mérés, vagy lehet több változós pl az időjárási adatok, páratartalom, hőmérséklet, barométer alapján következtetünk a várható időjárásra. Az idősor lehet stacionárius és nem stacionárius is. A nem stacionáriusra jellemzők hogy a variancia, átlag változnak az idő előre haladása során, vannak benne trendek, ciklusok. A stacionárius esetben nincsenek trendek, ciklusok, a variancia konstans az időtől független. Természetesen a stacionaritás nem ilyen fekete és fehér. létezik erős és gyenge stacionaritás is. Idősor előrejelzés nagyon sok helyen alkalmaznak. Időjárást, autóforgalmat, népességet, tőzsdét, fogyasztást is jeleznek előre. Nekem itt lesz egyszer egy tőzsdeszerű záróár és egy fogyasztási komponens.

## Trend elemzési modellek

## Regresszió

### Lineáris regresszió

A lineáris regresszió egyike a leggyakrabban alkalmazott statisztikai eljárásoknak, a trendelemzés alapja. Ha több folytonos változó lineáris kapcsolatban van egymással, akkor az egyik csoport segítségével (magyarázó változók) előre jelezhetjük a másik csoport értékét (eredmény változók). Szükségünk van a függő és független változó kiválasztására, de ez nem jelent oksági kapcsolatot! [9]

Az összefüggés segítheti a megértését a kapcsolatnak és legfőképp releváns előrejelzéseink lehetnek. Sajnos ez nem ilyen egyszerű.

Nagyon jó kiindulási alap, hiszen általa meg tudjuk határozni, hogy a változók függetlenek vagy összefüggnek-e, ha igen menyire vannak hatással egymásra, milyen a kapcsolat erősége, perdikció. Ezen okokból nagyon gyakran lehet vele több változós tőzsdei előre jelzésekben találkozni. [10] A trend meghatározás gyakori módszere.

Kétváltozós eset:

Y az eredmény változó

X a magyarázó változó

és a regressziós együtthatók, míg az a véletlen változó.

A lehető legkisebb hibájú becslés a cél. A hibáról feltételezzük, hogy független a magyarázó változótól és átlaga nulla. A becslés csak tökéletes kapcsolat esetén lenne hibamentes.

#### Becslési módszerek pl.:

* Legkisebb négyzetek módszere
* Momentumok módszere
* Legnagyobb valószínűségek módszere

#### Legkisebb négyzetek módszere

A **legkisebb négyzetek módszere** a mérések matematikai feldolgozásában használt eljárás. Nevét arról kapta, hogy az eltérések négyzetösszegét igyekszik minimalizálni.

A módszer érzékeny a nagyon kilógó adatokra. Egy kilógó adat az egész eljárás eredményét megváltoztathatja, hamis képet adva az adatsorról. Különböző statisztikai tesztekkel szűrik az adatsort, hogy ne maradjanak benne mérési hibák. A kilógó adatokat elhagyják, vagy a kívülállókra kevésbé érzékeny módszerekkel alternatív becsléseket végeznek. Ilyen például a súlyozott regresszió, amiben a kívülálló adatok súlyát, és ezzel befolyását is csökkentik. [12]

#### Elaszticitás

A rugalmasság mérőszáma. Azt fejezi ki, hogy az x magyarázó változó 1% változása hány %-os változást okoz az eredmény változóban.

## Mozgó átlagolás

A mozgó átlagolás elméleti összefoglalója [13] honlap alapján készült.

A mozgóátlagok számítása az idősorok hosszabb távú elemzésének legegyszerűbb módja. Csak annyit tűzünk ki célul, hogy átlagolással kiszűrjük a durva, egészen rövid távú ingadozásokat. Végrehajtása matematikailag igen egyszerű, jól alkalmazkodik az idősor jelleméhez. Hátránya hogy az idősor megrövidül, és jól kell, megválasztani az átlagolandó tagok számát különben torzít. Simítja az idősort, de az extrém értékek erősen befolyásolják. Mindig páratlan számú elemet tudunk átlagolni, mert az adott tagnak az átlagát úgy számítjuk, hogy vesszük az adatott tagot és előtte egy utána ugyan annyi számú további tagot. Ebből pedig már egyértelműen következik, hogy a mozgó átlagoknak van egy igen lényeges tulajdonsága, mégpedig az, hogy nem lehet minden egyes elemhez mozgóátlagot számítani, így a megfigyelt idősor eleje és vége elvész. Gyakran használják negyed éves trendek kiszámítására a devizapiacon vagy a tőzsdén. Egyszerű, nem igényel nagy gazdasági vagy matematikai tudást az alkalmazása. Készítenek egy rövid távú mozgóátlagot és egy hosszú távút, ha rövid távú felülről lefelé halad át a hosszú távún akkor várhatón csökkenő trend fog következni, így érdemes eladni, ha a rövid távú alulról metszik a hosszú távút akkor emelkedésre számítunk, így veszünk.

A mozgóátlagnak a 3 legismertebb fajtája a következő: egyszerű (vagy aritmetikai), exponenciális, és súlyozott. Módszerek közötti különbség a súlyozásban van, az egyszerű mozgóátlag minden egyes elemet egyforma súllyal vesz figyelembe, az exponenciális és a súlyozott pedig a frissebb adatokat nagyobb súllyal értékeli.

### **Egyszerű mozgóátlag**

Az adatsor egyszerű számtani átlaga, ahol azonban -amint azt már fentebb említettünk-, nem szabad elfelejteni, hogy az egyszerű mozgóátlag késve követi a folyamatokat, így a trend megváltozását is késve jelzi. Mivel miden adatot egyforma súllyal vesz figyelembe, nem számol azzal a ténnyel, hogy a frissebb adatok jelentősége nagyobb.

A napok megválasztását illetően annyit érdemes még megjegyezni, hogy kevesebb nap megválasztása esetén a mozgóátlag gyorsabban és érzékenyebben reagál a változásokra

A tőzsdei gyakorlatban azonban a mozgóátlagokat nem úgy számítják, hogy a kiválasztott adat környezetében végzik az átlagolást, hanem a kiválasztott adat előtti adatokra, így nem veszik el annyi darab mozgóátlagunk, mint a választott periódus fele. Az átlag számítása tehát nem centrikus, hanem visszatekintő.

### Exponenciális simítás,

Egyszerű modell. Nem feltételezünk az idősoros adatokban sem trend, sem szezonális hatást. Ez is egy fajta mozgó átlagolás azzal a különbséggel, hogy egy adott pont exponenciális simításának értékéhez elegendő a múltban közvetlenül előtte levő értékeket ismerni. Minden korábbi pont visszafelé haladva egyre kisebb súllyal számít. A súly értéke 0 és 1 között lehet. Az 1hez közeli értékek nagy súlyt kapnak az aktuális pont kiszámításában, kevésbé simítanak, míg 0 közeli súlyok erős simítást végeznek.

#### Egyszeres exponenciális simítás

Az exponenciális simítás módszer alapváltozata. Feltételezi, hogy a megfigyelt érték egy állandó körül ingadozik. A simítást α -val jelöljük. Az alfa azt mondja ki hogy mennyire szeretnénk simítani, azaz mennyire vesszük figyelembe az elkövetet hibát. Az alfa helyes megválasztása kulcsfontosságú.

Az egyszeres exponenciális simítás egyenlete: ahol:

0 ≤ α ≥ 1

a t időszakra vonatkozó trend szerinti érték

a t időszakra vonatkozó tényleges megfigyelés

Az egyszeres exponenciális simítás csak egy időszakra ad érdemi előrejelzést. Feltétele hogy az idősorban ne legyen tartós tendencia. Nagyobb alfa esetén jobban követi a tényleges adatokat, míg kisebb alfával hosszabb távú ingadozást lehet meghatározni. Ezért mikor egy idősor feltehetőleg lineáris trendet követ, nem célszerű ezt a módszert használni.

#### Brown féle kettős exponenciális simítás

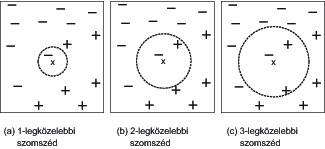
A kétszeres simítás nem más, mint az egyszeresen simított sor ismételt egyszeres simítása, így figyelembe tudjuk venni a trend értékét is. A fentebbi képletben a simított és az előre jelzett értékek megegyeztek, most viszont a kétszeres simítás miatt nem ugyan azok. S lesz a simított érték és felső indexbe teszem a simítás számosságát. Ennek megfelelően az egyszeres simítás képlete:

A kétszeres simítás egyenlete pedig:

A kettős simításnak hála alfa = 1 esetében a torzítás teljesen megszűnik, viszont ez a módszer add teljesen jó megoldást a kis alfák túl simítására. Ezzel a módszerrel tetszőleges hosszú és számú időszakra tudunk előrejelzést készíteni. Már kevés adattal is használható.

## Legközelebbi szomszéd

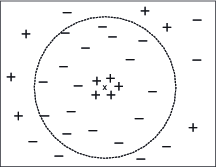
A legközelebbi szomszédok használatának indoklását legjobban a következő mondás szemlélteti: *" Ha valami úgy totyog, mint egy kacsa, úgy hápog, mint egy kacsa és úgy néz ki, mint egy kacsa, akkor az valószínűleg egy kacsa.''* A legközelebbi szomszéd osztályozó minden egyes esetet egy adatpontként reprezentál egy http://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_adatbanyaszat/math/eq_1435.png -dimenziós térben, ahol http://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_adatbanyaszat/math/eq_1436.png az attribútumok száma. Egy adott teszteset esetén a meghatározott szomszédsági mértékek valamelyikével kiszámítjuk annak közelségét a tanulóhalmaz összes többi adatpontjához. Egy adott http://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_adatbanyaszat/math/eq_1437.png eset http://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_adatbanyaszat/math/eq_1438.png -legközelebbi szomszédja azt a http://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_adatbanyaszat/math/eq_1439.png pontot jelenti, amelyek a legközelebb vannak http://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_adatbanyaszat/math/eq_1440.png -hez.



Az ábra a körök középpontjában lévő adatpont 1-, 2- és 3- legközelebbi szomszédját szemlélteti. Egy adatpontot a szomszédainak osztálycímkéje alapján osztályozunk. Abban az esetben, ha a szomszédoknak egynél több címkéje van, az adatpontot a legközelebbi szomszédok többségi osztályához rendeljük hozzá. Az ábrán az adatpont 1-legközelebbi szomszédja egy negatív eset. Ezért az adatpontot a negatív osztályhoz rendeljük hozzá. Ha három legközelebbi szomszéd van az ábrán látható módon, akkor a szomszédság két pozitív és egy negatív esetet tartalmaz. A többségi szavazási sémával az adatpontot a pozitív osztályhoz rendeljük hozzá. Holtverseny esetén az adatpont osztályozásához véletlenszerűen választhatjuk valamelyik osztályt.

A fent leírtak http://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_adatbanyaszat/math/eq_1441.png helyes megválasztásának fontosságát hangsúlyozzák. Ha http://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_adatbanyaszat/math/eq_1442.png túl kicsi, akkor a legközelebbi szomszéd osztályozó a tanulóadatokban jelenlevő zaj miatt hajlamos lehet a túlillesztésre. Másrészt viszont, ha http://www.tankonyvtar.hu/hu/tartalom/tamop425/0046_adatbanyaszat/math/eq_1443.pngtúl nagy, akkor a legközelebbi szomszéd osztályozó rosszul osztályozhatja a tesztpéldányt, mivel a legközelebbi szomszédok listája a szomszédságtól messzi adatpontokat is tartalmazhat

k -legközelebbi szomszéd osztályozás nagy k esetén



## ARMA modell

Autoregresszív és mozgótálag modell

## Modellek összehasonlítása:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Módszer | Főbb jellemzői | Előnyei | Hátrányai | Milyen időtávra |
| Lineáris regresszió |  | Egyszerű | Pontatlan, ha sztochasztikus az adat |  |
| Mozgó átlagolás | Az idősor dinamikus átlagát állítja elő | Egyszerű  Kevés adattal is működik | Pontatlan | maximum negyedév |
| Exponenciális kiegyenlítés | Különböző súllyal veszi a vizsgált időszak résztendenciáit | Rövidtávon megbízható | Nehéz meghatározni a súlyokat | Rövidtáv, középtáv |
| Legközelebbi szomszéd | Hasonlóság vizsgálat | Már 2 nap után tud működni | Hasonlóság mértékét nehéz definiálni, a jó működéshez sok adat kell |  |

Lineáris regresszió vizsgálata az adatainkon

# Neurális hálók

Az informatikában a 2010es évekre a GPU-k és a CPU-k exponenciális fejlődése lehetővé tette, hogy eddig már ismert, de túl elterjed matematikai eljárásokat, elveket elővegyenek és alkalmazzanak. Az egyik ilyen fejlődési irány a gépi tanulás, neuron hálók világa. Lassan nem lesz olyan ágazat, ami ne használná valamire a neuron hálót. Képet, hangot ismertetünk fel vele, jó osztályozásra, előre jelzésre. Olyan nagy cégek, mint az Nvidia vagy az Oracel ölnek időt és pénzt a neuron hálók fejlesztésébe. Egyre több programozási nyelvhez érhető el neuron hálós kiegészítés.

Maguk a neurális hálók nem új keletűek már 1958-ban Rosenblatt megalkotta a perceptron elméletét ráadásul neuronhálókkal kapcsolatos kutatások a 90-es évek kezdete óta már Magyarországon is folynak. A neurális hálózatok olyan, számítási feladatok megoldására létrejött párhuzamos feldolgozást végző, adaptív eszközök, melyek eredete a biológiai rendszerektől származtatható. A neurális hálózatok számos feladat megoldásánál nemcsak alkalmasnak, hanem alapvetően jobbnak is bizonyulnak, mint a hagyományos algoritmikus számítási rendszerek. Ilyen feladatok tipikusan a különféle felismerési problémák, kezdve a viszonylag egyszerű nyomtatott számok és karakterek felismerésétől a jóval bonyolultabb kézírás-, kép- és egyéb alakzatfelismerésekig.

A felismerési feladatokon kívül sok egyéb olyan problémával találkozunk, ahol jelenleg nem ismert algoritmikus megoldás, vagy ha van is ilyen, az annyira bonyolult és/vagy olyan sok műveletet igényel, hogy reálisan elfogadható idő alatt a mai legnagyobb teljesítményű számítógépekkel sem oldható meg.

E feladatok egy részénél − ilyenek pl. komplex ipari, gazdasági vagy pénzügyi folyamatok időbeli viselkedésének előrejelzése − a megoldás nehézsége általában abból ered, hogy nem rendelkezünk azzal a tudással, amely az algoritmikus megoldáshoz szükséges lenne. Alapvetően kétféle tudásra lenne szükségünk: egyrészt a folyamatok mögött meghúzódó fizikai, közgazdasági, stb. törvényszerűségeket kellene ismernünk, másrészt a törvényszerűségek ismeretében is csak akkor tudnánk a feladatot megoldani, ha a folyamatokat létrehozó rendszerek pillanatnyi állapotát és e rendszerekre ható összes vagy legalább minden lényeges hatást, a környezeti feltételeket is ismernénk. A feladatokról ugyanakkor más formában − adatokban megtestesülve − rendelkezésünkre áll tudás, amelyet ha fel tudunk használni, a feladat valamilyen megoldásához eljuthatunk. Egy adott az adatok által hordozott tudás sohasem teljes. Mégis a neurális hálók képesek az adatokból nyert ismeretek általánosítására: olyan szituációkban is jó választ adnak, melyek az adatok között nem találhatók meg. Egyes mesterséges neurális hálók is rendelkeznek ezzel a képességgel. Ez azt jelenti, hogy hiányos, esetleg pontatlan ismereteket hordozó, legtöbbször zajos adatokból is kinyerhető általános tudás. Ez a tudás azonban a feladat "tökéletes" megoldását rendszerint nem teszi lehetővé, viszont "jó" megoldás elérését biztosítja.

# Irodalom jegyzék:

1. <http://www.bibl.u-szeged.hu/ha/gazd/gazdasag/kl.html> kilingrendszer fogalma

1. <https://www.hupx.hu/hu/Piacosszekapcsolas/piacosszekapcsolastort/Lapok/default.aspx> 4M fogalma + története
2. [https://www.MAVIR.hu](https://www.mavir.hu)

1. <https://www.mvmpartner.hu/hu-HU/Szolgaltatasok/Villamos-Energia/Tudastar/FogalomTar> Mavir fogalma
2. [https://www.MAVIR.hu/documents/10258/107815/Sz%C3%A9lkihaszn%C3%A1lts%C3%A1g+tanulm%C3%A1ny+2010.pdf/153d2d78-1c6f-4d54-858e-5bc46f56c352](https://www.mavir.hu/documents/10258/107815/Sz%C3%A9lkihaszn%C3%A1lts%C3%A1g+tanulm%C3%A1ny+2010.pdf/153d2d78-1c6f-4d54-858e-5bc46f56c352)

1. [http://www.MAVIR.hu/c/document\_library/get\_file?uuid=81fd9f45-12cf-44e3-ad74-9492504a42ef&groupId=10258](http://www.mavir.hu/c/document_library/get_file?uuid=81fd9f45-12cf-44e3-ad74-9492504a42ef&groupId=10258) a határmetszék definiciója innen van
2. <http://tozsdebarat.oldalunk.hu/site.php?sd=tozsdebarat&page=oeoxtYiFA0>
3. <http://www.cs.bme.hu/nagyadat/>
4. <https://www.mateking.hu/statisztika-2/regresszioszamitas/>
5. <http://www.cs.bme.hu/nagyadat/bartok-ferenc.pdf>
6. <http://psycho.unideb.hu/munkatarsak/balazs_katalin/stat1/stat1ora4.pdf>
7. <https://hu.wikipedia.org/wiki/Legkisebb_n%C3%A9gyzetek_m%C3%B3dszere>
8. <http://www.portfolio.hu/vallalatok/technikai-elemzes/technikai-elemzes-a-mozgoatlag-hasznalata.17774.html>
9. <https://www.elemzeskozpont.hu/mozgoatlagok>
10. <http://ilias.gdf.hu/data/ilias-ha/lm_data/lm_9370/index.html>
11. <http://gtk.uni-miskolc.hu/files/8449/Exponenciális+kisimítás.pptx>
12. <http://statisztikus.hu/fuggelek/>
13. <http://elib.kkf.hu/okt_publ/szf_11_02.pdf>
14. <http://www.portfolio.hu/vallalatok/technikai-elemzes/technikai-elemzes-a-mozgoatlag-hasznalata.17774.html>
15. **A predictive pan-European economic and production dispatch model for the energy transition in the electricity sector, PowerTech, 2017 IEEE Manchester, 18-22 June, Laurent Pagnier, Philippe Jacquod**  
    <http://ieeexplore.ieee.org/document/7980982/>
16. Modeling and simulation inspired by quantum methods of the Polish Electricity Stock Exchange**,** Progress in Applied Electrical Engineering (PAEE), 2017 **25-30 June,** Jerzy Tchorzewski, Dariusz Rucinski<http://ieeexplore.ieee.org/document/8008983/>
17. ANN approach for predicting economic trends based on electric energy consumption during natural disaster period, Knowledge, Information and Creativity Support Systems (KICSS), 2016 11th International Conference on, 10-12 Nov**,** Akanit Kwangkaew, Virach Sornlertlamvanich, Itsuo Kumazawa  
    <http://ieeexplore.ieee.org/document/7951405/>
18. Neurális hálózatok, Altrichter Márta, Horváth Gábor, Pataki Béla, Strausz György, Takács Gábor, Valyon József, 2006 Panem Könyvkiadó Kft., <https://mialmanach.mit.bme.hu/neuralis/pr02>