

Időjárás-jelentések összehasonlítása

Rácsik Patrícia

Témavezető: Machalik Károly

2020

Pannon Egyetem

Műszaki Informatikai Kar

Rendszer- Számítástudományi Tanszék

Programtervező Informatikus BSc

Nyilatkozat

Köszönetnyilvánítás

Tartalmi összefoglaló

Az elmúlt évtized hatalmas technikai fejlődésének köszönhetően ma már szinte mindenki hozzáfér olyan felületekhez, melyek az ember mindennapi életének megkönnyítésére szolgálnak. Ezek közül nagy népszerűségnek örvendenek az időjárás előrejelző oldalak. De gondolkoztunk-e már azon, hogy ezek az információk mennyire hasonlítanak vagy különböznek egymástól? A hétköznapi felhasználó azt várná el, hogy ebben a témakörben különböző felületeken, ugyan eltérő formában, de hasonló adatok birtokába juthat. Sajnos a tapasztalat ettől jelentős mértékben eltér és egyelőre ezen felületek tartalmának összehasonlítására sincs megoldás. Szakdolgozatomban a fent felvetett problémára keresem a megoldást. A dolgozatban különböző statisztikai és numerikus módszerekkel vizsgálom az eltérő időjárási adatokat. Szó lesz alkalmazásprogramozási interfészekről (API), az általuk biztosított adatok feldolgozásáról, adatbázis tervezéséről, adatok elemzéséről, vizualizálásról.

Abstract

Tartalomjegyzék

1. Bevezetés	7
2. Időjárás alapjai	9
2.1. Fontos tényezők	9
2.2. Meteorológia	12
2.3. Emberre való hatása	15
3. Használt technológiák	17
3.1. Alkalmazási interfészek	17
3.2. Python programnyelv	19
3.3. Relációs Adatbázisok és Microsoft Sql Server . . .	21
4. Adatbázis tervezése	22
5. Adatgyűjtés Python segítségével	22
5.1. API hívások, API-k amiket használok	22
5.2. JSON Feldolgozás	22
5.3. Adatbázis és Python kapcsolata	22

1. Bevezetés

Témaválasztásom fő mozgató rúgója az volt, hogy megismerjen azon jelenleg rendelkezésre álló adatelemzésre, összehasonlításra és vizualizálásra alkalmas módszereket melyek, alkalmasak egy olyan kérdés megoldására, mint például az időjárási adatok összehasonlítása. Talán az időjárás, mint fogalom elsőre nem hangzik túlságosan érdekesnek a hallgatóság számára, viszont a téma mégis rengetek, dolgot magába foglal, lehetőség nyílik új ismeretek elsajátítására nem csak az informatika világában, de a matematika területén is. Valamint a téma hatalmas irodalmi terjedelemmel rendelkezik, rengetek platform ingyenes illetve korlátozott adathozzáférést biztosít, melyek mind segítik az elemző munkámat.

Az időjárás az élet minden területén megtalálható. Nap, mint nap találkozunk vele, de pontos tájékoztatást ritkán tudhatunk a magunkévá. Véleményem szerint az átlagember nem is mérlegeli, hogy a különböző időjárás előrejelző platformok mennyire eltérő adatokkal szolgálnak. Ha jobban beleássuk magunkat, talán azt mondhatjuk nincs két egyforma tájékoztatás. Pedig egységes és megbízható időjárás előrejelzés minden bizonnyal nagymértékben megkönnyítené az életünk. Dolgozatban erre is keresem a megoldást, melyik az általam vizsgált legmegbízhatóbb platform és nagyjából mi a pontos időjárási viszonyosság, amire készülnie kell

az embernek.

Fontos megemlíteni, hogy a dolgozatomban nem csak e probléma megoldása motivált, hanem szakmai kíváncsiság is vezérelt. Szerettem volna komolyan megismerni az adattudományok világát, több ismertetet szerezni az adatbázisok terén, valamint elsajátítani a Python programozási nyelv szépségét, előnyeit.

A dolgozatot két fő részre osztottam. Az első részében részletezem a dolgozathoz szükséges háttér-információkat, szakmai irodalmat valamint a későbbiekben használt technológiákat. A következő részben pedig az általam készült szoftver készülők bemutatni, elemezni a feltárt eredményeket.

A dolgozat írása során főként a magyar és külföldi szakirodalmat, internetes cikkeket valamint tanulmányaim során elsajátított ismereteimet használtam fel.

2. Időjárás alapjai

Az időjárás a napi hőmérsékletre és csapadék aktivitásra utal, aktivitására illetve változására utal vagy éppen írja le mint például meleg vagy hideg, nedves vagy száraz, szeles, napos vagy esős az idő. Ezek a fizikai változások az atmoszférában az az a légkörben történő jelenségek következménye. Ezek az események általában a légkör legalacsonyabb szintjén a troposzférába, a sztratoszféra alatt következnek be. Nagyonbár az időjárás a Föld időjárására utal. Éghajlat kifejezés a légköri viszonyok hosszabb ideig történő átlagolására reflektál. Légkörünkben több egységet különböztünk meg, ezek rendre a következők:

- Troposzféra
- Sztratoszféra
- Mezoszféra
- Termoszféra
- Exoszféra

2.1. Fontos tényezők

Az időjárást rengeteg tényező befolyásolja, ezek közül a legfontosabbak: légnyomás, páratartalom, szélirány, szélereősség, felhőzet

valamint a hőmérséklet.

Nyomás egy felületi egységre ható nyomóerő, fizikai mennyiség amelyet a nyomás felületre ható erő és felület hányadosa által értelmeziünk. Légnyomás esetén a nyomóerőt a légkör egy adott pontja felett elhelyezkedő levegőoszlop súlya határozza meg. Meteorológiába ennek mértékét hektopascalban (hPa) definiálják. A légkörben emelkedő irányba haladva közelítőleg exponenciálisan csökken. Akkor beszélhetnénk teljes exponenciális csökkenésről, ha a légkör hőmérséklete változatlan maradna.

Egy hidegfront esetén hideg levegő érkezik a fennálló meleghez, a terjedelmesebb tömeg miatt a meleg légtömeg alá nyomja magát, ennek eredményeképp a légnyomás először csökken aztán pedig emelkedik. A jelenséget csapadék és hőmérséklet lehűlés követi. Ugyan így történik melegfront esetén is, ebben az esetben a meleg levegő a hideg légtömeg alá furakodik, a légnyomás permanensen csökken de a hidegfronttal ellentétben nem emelkedik hanem marad az adott szinten. Okklúziós frontról beszélhetünk ha a hideg és a meleg egyszerre érkezik meg. Általánosságában kijelenthetjük, hogy a melegfrontot utoléri egy gyorsabban mozgó hideg, s ennek hatásai érvényesülnek az időjárásban.

Pára akkor jön létre ha a Nap felmelegíti a vizet, ennek következtében pedig a levegőbe kerül vízgőz formájában. Ennek

mennyisége attól függ , hogy a víz melegedésének mértéke mekkora.. Ezzel párhuzamosan a légkör párapacitása is növekszik. Páratartalom definiálására két lehetőséget különböztetünk meg, abszolút vízgőztartalom valamint a relatív vízgőz tartalom. Az abszolút vízgőztartalom segítségével kimutathatjuk ,hogy $1m^3$ levegőben milyen mértékű a grammban vett vízgőz tartalom. Ezzel ellentétben a relatív vízgőz tartalom azt jelenti , hogy a levegőben lévő vízpára az adott hőmérsékleten milyen százalékos értéket mutat, a vízgőzzel teljesen telített levegő víztartalmához mérten.

A levegő mely légkörünk alkotó eleme, szélnek nevezzük ha közelítő vízszintes irányú áramlását végezz. A jelenség a helyi nyomáskülönbségeknek tudható be. Sebességét méter/másodpercben, erősségét Beaufort-skálán osztályozzák. A föld forgásából adódóan három egymáshoz kapcsoló szélrendszer különböztetünk meg:

- Passzátszél
- Nyugati szél
- Keleties sarki szél

A szél irányát az határozza meg mely égtáj felől érkezik. Tehát keleti szélről beszélhetünk, ha a szél keleti irányból nyugat felé fúj, ami az jelenti, hogy kelet felé fordulva szembe érkezik. Sok

esetben fokkal is jellemzik, 0° északi, 90° keleti, 180° déli valamint 270° esetén nyugati szélnek nevezzük ezeket. Meteorológiában 10 fokos pontosságot használnak. Levegőmozgásának sebessége alatt, magát a szél sebességét értelmezzük. A meteorológusok számára Beaufort-skála szolgál a szél sebességének jellemzéséhez. Szélfajták: Ciklon, Hurrikán, Tájfun, Monszun, Tornádó, Orkán.

Levegőbe a felszálló vízpárából parányi vízcseppek és jégkristályok halmaza alakul ki, ezek a légkörünkben találhatóak, mindennapokban felhőnek nevezzük. Ezek eredményeképp csapadék, eső vagy akár jégeső is keletkezhet, a talaj közelében is találkozhatunk vele, köd formájában. Diszperz rendszer azaz egymástól független részecskékből áll, amely a fény útját is elejét veszi. Magasságuk és kiterjedésük alapján osztályozzuk őket. A felhők tipizálásánál a következő tulajdonságokat vesszük figyelembe: méret, forma, szerkezet, textúra, fényhatás, szín, a felhőalap magassága. Forma szerint 14 felhőtípust különböztetünk meg, anyaguk szerint pedig víz-, jég és vegyes halmazállapotú felhőket.

2.2. Meteorológia



A meteorológia vagy ha régies nevét használjuk légtünemény-

tan, előrejelzésekkel és időjárási folyamatokkal foglalkozik. Időjárás megfigyelése már az ókorban elkezdődött, első feljegyzések már mezopotámiai kultúrából származnak, de találhatunk ókori görög-római és kínai annotáció is. Maga a kifejezés Arisztotelész Meteorológia című művéből származik. Galileo Galilei az elsők között készített termoszpókot, melynek segítségével bizonyította a hőmérséklet lehetőségét. Saját termoszkopja ugyanis jelezte a meleg levegő tágulásának jelenségét. Blaise Pascal fedezte fel, hogyha egyre csak emelkedünk a tengerszinttől akkor a légnyomás csökken, amiből következtethetünk arra , hogy a légkör felett vákuum van. Anders Celsius nevéhez fűződik az általunk is mindennap használt hőmérőbeosztás. Robert Hooke, Edmun Halley, George Hadley, Benjamin Franklin, Luke Howard és még rengetek más embert hozzájárult ahhoz, hogy napjainkban pontos méréseket készíthessünk.

Numerikus időjárás előrejelzés a 20. század elején vált elérhetővé, a légkör dinamika megismerésének és fejlődésének eredményeképp. Lewis Fry Richardson Időjárás-előrejelzés numerikus eljárásokkal című könyvében olvashatunk arról, hogy rövidebb időszakok előrejelzésére numerikus megoldás található, viszont csak fontos légköri áramlat elhanyagolásával lehetséges. A számításokhoz létfontosságú számok túlságosan nagyoknak bizonyultak egé-

szen az első számítógépek megjelenéséig. Az 1950-es évek elején a számítógépes számítások elérhetővé váltak. Az első időjárás-előrejelzések a barotropikus modellekből származnak, amelyekből sikeresen meg lehet jósolni a középmagas légkörben elhelyezkedő nagy mozgású Rossby hullámokat, amik az alacsony, illetve a magas légköri nyomást rajzolják ki.

Műholdas megfigyelések lehetőségének megkezdését az 1960-as években vált elérhetővé, TIROS-1 időjárás-előrejelző műhold ekkor került fellövésre. Segítségével rendkívül gyorsan, mondhatni pillanatok alatt adatokat juttattak a világ egyik pontjából a másikba. Az időjárási műholdak és a különböző magasságokban keringő általános célú megfigyelőműholdakkal elengedhetetlen eszközévé váltak. Napjainkban azon éghajlati modelleket fejlesztik, amelyek összeegyeztethetőek a régebbi időjárás-előrejelző rendszerekkel. Ezen éghajlati modellek segítségével megfigyelhetőek a hosszan tartó éghajlatváltozások.

Az időjárás előrejelzéséhez különböző eljárások léteznek:

- Változatlanság
- Tendencia
- Éghajlattan
- Analóg módszer
- Numerikus időjárás-előrejelzés

Változatlanság szerint a kondíciók nem változnak, hosszútávú előrejelzés esetén leghasznosabb módszer. Tendencia alatt értjük a frontok sebességének és irányának definiálását, valamint a magas és az alacsony légnyomással bíró lokációk, felhős és csapadékos területeket meghatározását is. Éghajlattan a tapasztalati adatok felhasználását jelenti az adott nap meghatározásához. Összetett eljárásnak mondhatjuk az analóg módszert mely régi adathalmazokból hasonló időjárási körülményeket mutat. Numerikus időjárás-előrejelzés számítógépek segítségével magas számú változók használatával, légköri modellt generál. Legsikeresebb és legszélesebben használt módszer.

2.3. Emberre való hatása

Manapság a populáció fele időjárás-érzékeny. Az időjárás-érzékenység

akár társadalmi betegségnek is mondható, ha természetes környezeti hatásoktól rendszeresen elszigetelődünk és ezáltal csökken a szervezeti teljesítőképessége, valamint megnő az érzékenységünk az időjárással szemben. Minél korosabb az adott egyén annál jobban reflektál az időjárás-változásokra. Mivel idővel lassabban játszódnak le a belső folyamatok, testünk nem tud oly módon igazodni mint korábban, így időjárás-érzékennyé válunk. Tünetek közé tartok az: idegesség, alvási zavarok, kimerültség de akár a depresszió, szív- és keringési zavarok is. Kisgyermek és csecsemők szintén reagálnak a változásokra. Esetükben csökken a koncentrációképesség valamint a felfogóképesség, kimerültebbé is válhatnak. Kutatások szerint a nők gyakrabban és jobban érzékenyebbek a szélsőséges időjárás változásokra mint férfi társaik. Ez a különbség már fiatalabb, kisiskolás korban is felmerül.

Az időjárás-érzékenység a földrajzi elhelyezkedésen is múlik. Példaként említeném, hogy a közép-európai népek rendkívül érzékenyek, mivel a közepes szélességeken élők, átlagosan akár hat naponta ki vannak téve frontátvonulásnak. Vizsgálatok az mutatják, hogy mi magyarok általában a melegfrontra vagyunk érzékenyek ellenben a franciák vagy az angolok a hidegre.

3. Használt technológiák

A szakdolgozat kezdeti fázisában alaposan átgondoltam, valamint mérlegeltem, hogy melyek azok technológiák amik elengedhetetlenek lehetnek a megvalósításhoz. Tanulmányaim során is találkoztam már nagy részükkel, viszont az internet segítségével és tájékozódtam, továbbá kikértem tanáraim és az egyetem munkatársainak tanácsát is. A dolgozat elkészítéséhez első sorban szükségem volt egy adatbázisra, már általam is jól ismert és használt Microsoft Sql Server-re esett a választásom. Az adatok begyűjtéséhez és vizualizálásához, elemzéséhez a Python programozási nyelv és különböző könyvtárainak lehetőségeivel éltem. Fontos még megemlítenem azt is, hogy az adatok begyűjtése alkalmazásprogramozási interfészek azaz API hívások segítségével történt. A következőben ezeket szeretném részletesen bemutatni, valamint szemléltetni melyek azok az okok, ami miatt az említett technológiák mellett döntöttem.



3.1. Alkalmazási interfészek

Angol nyelven application programming interface kifejezésként ismert, rövidítése API. Ha egyszerűen próbáljuk megmagyarázni az API tulajdonképpen

egy olyan szoftverközvetítő közeg, ahol két applikáció képes egymással kommunikálni. Máshogy fogalmazva, az API egy üzenetküldő ami az egyik oldal kérését továbbítja a másik oldal felé, majd ennek válaszát visszajuttatja a kezdeményezőnek. Ha egy fejlesztő egy alkalmazást készít, nem feltétlen a nulláról indul neki a feladatnak. Az API-k lehetővé teszik a fejlesztők számára, hogy ismétlődő, mégis összetett folyamatokat készítsenek egy rövidebb kódot felhasználva, és ezek a későbbiek folyamán újrafelhasználhatóvá válnak. Mivel az API-k kezelhetősége egyszerű és egységes, rendkívül hozzájárulnak ahhoz, hogy a mai alkalmazásfejlesztés gyors ütemével képesek legyen lépést tartani a fejlesztők, továbbá az alkalmazás működési sebességére is pozitív hatással van az, hogy az API kérések válaszideje szűk határidőn belül megvalósul.

Számos alkalmazási interfészt különböztetünk meg, mint például Java API vagy alkalmazásokon belüli interfészek, melyek lehetővé teszik az egyik objektumok közötti kommunikáció a Java nyelvben. A programközpontú API-k között még találkozhatunk olyan web alapú alkalmazási interfészekkel mint például a Simple Object Access Protocol(röviden SOAP) melynek üzenetek XML alapúak, továbbá a távoli eljáráshívás, angolul Remote Procedure Call ismert. Fontos megemlíteni a REST azaz Representational State Transfer API-t is melyek talán napjainkban a legelterjedtebb és legismertebb API. Nagyjából 15 000 publikus, bárki számára elérhető API elérhető a programozók számára, ezen felül rengetek magán kézben lévő, amelyeket a vállalatok saját belső és külső kapacitásuk bővítésére használnak.

De miért is fontos ez számunkra és miben járul hozzá a megvalósításhoz? Ahhoz, hogy időjárást tudjuk elemezni szükségünk van nagy mennyiségű időjárási adatra. Ezeket különböző már említett nyilvánosan elérhető API-k segítségével könnyen birtokolhatjuk. Az adatokat API hívás segítségével

tudjuk elérni. Sokféle, különböző típusú kérés létezik. Az adatok elkérésére a leggyakrabban használt módszer a GET kérés. Mivel esetünkben csak adatok lekérésére létfontosságú, ezért a hangsúlyt GET kérések benyújtására fordítjuk. Amikor az kérést teszünk a felület felé, az API-tól kapott válasz válaszkód formájában érkezik meg, ez mutatja a művelet sikerességét. Kódok azért is fontosak mert megtudhatjuk mit rontottunk el egy hibás hívás esetén. Annak érdekében, hogy sikeres kérést nyújtsunk be, amikor API-kkal dolgozunk, fontos, hogy olvassa el a hozzátartozó dokumentációt. Egy adott kiszolgálón gyakran található több API is, ezeket az API-kat endpointoknak azaz végpontoknak nevezzük. Végpontok tartalmazzák az adatokat melyek számunkra fontosak lehetnek. Ezek legtöbb esetben JSON formátumban vannak megadva.

A JSON (JavaScript Object Notation) az API nyelve, adatszerkezetek kódolására alkalmas. Főként adatok továbbítására alkalmazzák szervezek és web alapú applikációk között. Ember számára is olvasható adatcserére tervezték és modern programnyelvek mellett is könnyen használható.

3.2. Python programnyelv

Python általános célú és maga szintű programozási nyelv. 1989 végén Guido van Rossum holland származású programozó kezdett el rajta dolgozni, majd 1991-ben hozta nyilvánosságra. Készítésekor a fő szempontok közé tartozott, hogy könnyen olvasható legyen és programozói munka megkönnyítése előtérbe kerüljön szemben a futási sebességgel. Támogatja a funkcionális, az objektumorientált, az imperatív és a procedurális programozási paradigmákat. Hasonlít a Perl, Ruby, Scheme nyelvekhez, valamint szigorú típusrendszer birtokában van. A legfrissebb PYPL Popularity of Programming Language rangsorolása szerint, a tíz legnépszerűbb programozási nyelv egyike. Fel-

használható asztali grafikus felhasználói felületek, webhelyek és webes alkalmazások fejlesztésére is. A programozási nyelv egyszerű szintaxisszabályai továbbá megkönnyítik a kódok olvashatóságát és az alkalmazás karbantarthatóságát. Számos oka van annak miért választottam a Python nyelvet a dolgozat alapjaként.

Első szempont az volt, hogy könnyen éltethető és olvasható legyen bárki számára. Python könnyen elsajátítható programnyelvek közé tartozik. Ez elsősorban az angol nyelvhez való hasonlóságának köszönhető. A nyelv szintaxisára jellemző hogy nagyon kevés a szabály és különleges eset. Nyugodtan kijelenthetjük, hogy a Pythonban a kódolásra fókuszál, és nem a nyelvi bonyolultságok helyezi előtérbe. A programozási nyelv másik vonzó aspektusa a hatékonysága és az olvashatósága.

Python táborát erősítette még az is, hogy Bross-Platform és nyílt forráskódú. Több mint húsz év telt el azóta, hogy a szóban forgó nyelv több platformon fut és nyílt forrású Legyen Linux, Windows vagy MacOS, a Python-ban írt kódok minden platformon működnek. Ezek mellett sokoldalú nyelv és platform: A Python továbbra is nagyon releváns, mivel bármilyen szoftverfejlesztésben felhasználható, legyen az lokális és felhőben kezelendő projekt. Támogatja objektum-orientált programozást is, de kisebb programok, tool-ok fejlesztését is alkalmas.

Data science és mesterséges intelligenciák legelterjedtebb nyelve napjainkban, mindkét témakör nagyban érinti a dolgozat témáját is. Python az adattudósok körében az egyik legelterjedtebb nyelv. Legyen szó szoftverfejlesztésről vagy marketingről, napjainkban minden munkaügyi szektor rengeteg adatot használ és hagy maga után, és ezek elemzése valamint megértése nagyban befolyásolja további működésüket, stratégiájukat. A 'Numpy' és a 'Panda' könyvtárak megjelenésével a Python kiemelkedő jelentőségűvé vált

az adatvilágban. A Python statisztikai, táblázatos és mátrix adatokat is kezel, ezeket „Matplotlib” és a „Seaborn” könyvtársak segítségével képes vizualizálni. Jelenünkben nem vitás, hogy a jövő informatikáját a mesterséges intelligencia vezetni fogja. A Python machine learning könyvtárai, például a Keras és a TensorFlow, lehetővé teszik a számunkra gépi tanulás fejlesztését is. Az olyan könyvtárak, mint például az OpenCV pedig elősegítik a számítógépes látást vagy a képfelismerést.

Valamint rengeteg grafikus felületet is támogat ezek között a projektben megjelenik a Python Kivy könyvtára, amely szintén cross-platform és nyílt forráskódú, valamint nagyon modern és letisztult külsőt kölcsönöz, abszolút megállja helyét napjainkban.

3.3. Relációs Adatbázisok és Microsoft Sql Server

Relációs adatbázisok olyan digitális adatbázisok melyek a relációs adatmodellen alapszik. A relációs adatbázisok kezeléséhez használt szoftvereket relációs adatbázis-kezelő rendszerek röviden RDBMS nevezzük. Számos relációs adatbázisrendszernek képes az SQL azaz Structured Query Language(magyarul lekérdezőnyelv) használatára, ez általa lehetőségünk nyílik lekérdezni és karbantartani az adatbázisok tartalmát. A táblák minden sora rendelkezik a saját egyedi azonosítóval, kulccsal. A táblák sorai összekapcsolhatók más táblák soraival, ha rendelkeznek úgy nevezett idegen kulccsal.

Habár a feladat nem követeli meg az Microsoft Sql Server használatát, mégis úgy láttam hosszútávon kifizetődőbb vele dolgozni.

elakadtam :((szívesen fogadokj tippeket miért is jobb az MSSQL, hirtelen csak ennyi van hogy: gyorsabb, több funkció talán, T-SQL

4. Adatbázis tervezése

Itt fogom bemutatni a táblákat stb

5. Adatgyűjtés Python segítségével

rövid ismertető

5.1. API hívások, API-k amiket használok

free oldalak bemutatása, dokumentáció és lehetőségek, api hívások pythonba

5.2. JSON Feldolgozás

Json feldolgozása python segítségével

5.3. Adatbázis és Python kapcsolata

milyen király hogy adatbázisba minden benne -> letudok kérni minden + tárolás