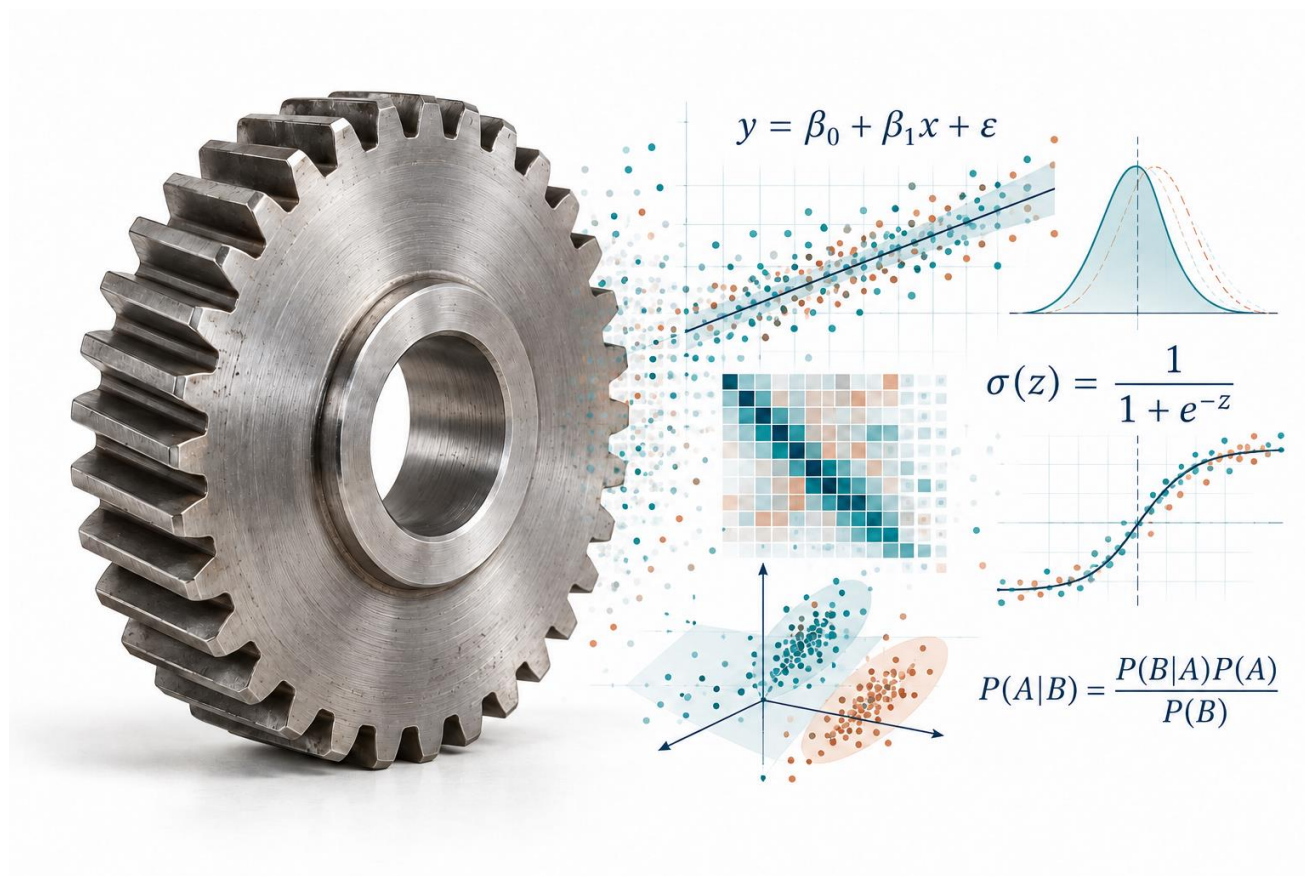


Kriston András

Alkalmazott data science mérnököknek

Egyetemi jegyzet



Szerzői jog és közreműködés

© 2026 Kriston András. Minden jog fenntartva.

A szerzői jogi fenntartás a szerző saját, szerzői jogi védelemben részesülő hozzájárulásaira vonatkozik. A jogszabályban megengedett felhasználások kivételével a jegyzet egészének vagy védett részeinek többszörözése, terjesztése, átdolgozása és nyilvános hozzáférhetővé tétele a jogosult előzetes engedélyéhez kötött.

A nyilatkozat nem korlátozza a jogszerű idézést és az egyéb törvényi szabad felhasználásokat. A más szerzőktől származó idézetek és egyéb átvett tartalmak jogai az eredeti jogosultakat illetik. A forrásmegjelölés önmagában nem jelent általános felhasználási engedélyt.

A jegyzet hivatkozása

Kriston, A. (2026). Alkalmazott data science mérnököknek.

A szerzőről

Kriston András data scientist, gépészmérnök. Bsc és Msc képzését a Miskolci Egyetem Gépészmérnöki és Informatikai Karán folytatta 2006 és 2012 között. A PhD fokozatot 2017-ben szerezte meg a finn Aalto University Gépészmérnöki Karán (Otaniemi, Espoo Finnország), ahol kutatási témája a gumi-jég tribológiai folyamatainak vizsgálata volt. Az adattudománnyal először mint digitális képfeldolgozással foglalkozott doktori kutatása alatt, majd későbbiekben ezzel a témában foglalkozó kutatóintézetben is dolgozott. A gyártási feladatokkal kapcsolatos adattudomány áll jelenleg a munkája középpontjában.

Mesterséges intelligencia közreműködése

A jegyzet szakmai kiindulópontját, mérnöki tapasztalatait, alapgondolatait és eredeti szövegrészeit Kriston András adta. A kidolgozás során az OpenAI ChatGPT eszközt használta szerkesztési és technikai segítségként.

A jegyzet szintetikusnak jelölt ábrái mesterségesen előállított adatokból készültek. Ezek oktatási szemléltetések, nem valós üzemi mérések.

A ChatGPT eszközként működött közre, nem társszerzőként vagy független szakmai lektorként.

2026. szeptember 30.

Alkalmazott data science mérnököknek

Mi is a data science?

Mi is valójában a data science, vagyis az adattudomány? Hol van a szerepe ma, egy olyan világban, ahol egyre gyakrabban úgy tűnik, hogy a mesterséges intelligencia varázsütésre képes megoldani szinte bármilyen problémát?

Milyen adatok állnak a mérnöki adattudomány fókuszában? Mi a gépi tanulás (machine learning), és hogyan alkalmazhatják a mérnökök és kutatók a saját szakterületükön? Mire használhatók az adatvezérelt modellek, és legalább ilyen fontos kérdés: mire nem?

Ez a kurzus és a hozzá kapcsolódó jegyzet elsősorban a műszaki területeken előforduló problémák megértésére és megoldására használható adattudományi módszerekkel, technikákkal és technológiákkal foglalkozik.

A statisztikai elemzésen és modellezésen alapuló módszerek segítségével többek között feltárhatjuk a folyamatokban rejlő összefüggéseket, azonosíthatjuk a problémák lehetséges okait, felismerhetjük a rendellenes működést, valamint előre jelezhetünk jövőbeli állapotokat. Mindezek célja végső soron az, hogy az adatokból olyan információt nyerjünk, amely jobb mérnöki döntések meghozatalát teszi lehetővé.

Az adattudomány tudományos szemléletének alapja, hogy eredményeit - modellek előrejelzéseit, illetve a modellek által feltárt összefüggéseket - kísérletileg is igyekszünk bizonyítani. Más tudományterületekhez hasonlóan (pl.: fizika, kémia, biológia), el kell mélyülni a folyamatok megfigyelésében, kapcsolatokat keresni a jelenségek között, kísérleteket tervezni és megvalósítani, majd a hipotéziseinket cáfolni vagy igazolni.

Ma már szinte minden nagyobb vállalatnál találkozhatunk adatelemzőkkel, adattudósokkal, adatmérnökökkel és különböző AI-hoz kapcsolódó munkakörökkel. A feladatkörök megnevezése és határai azonban rendkívül gyorsan változnak, ezért ebben a jegyzetben nem célok ezek részletes szétválasztása.

Sokkal fontosabb az együttműködés módja.

Az "adatos" szakemberek a műszaki terület szakértőivel közösen dolgozva próbálják megoldani a különböző problémákat, amelyek többek között gyártási, tervezési, folyamatoptimalizálási, minőségügyi területekről származhatnak.

Egy adattudós ismerheti a legmodernebb algoritmusokat, de ha nem érti, hogyan keletkezett egy mérési érték, mit jelent fizikailag egy változó, vagy hogyan működik maga a vizsgált folyamat, könnyen rossz következtetésre juthat, amely egy nem optimális megoldást eredményezhet.

Ugyanakkor a mesterséges intelligenciával támogatott programozási és adatelemzési eszközök fejlődésével a nem informatikai háttérű szakemberek is egyre összetettebb elemzéseket és fejlesztéseket képesek elvégezni. Persze, ahhoz, hogy valóban hatékonyan tudják használni a rendelkezésre álló eszközrendszereket, az adattudomány különböző területeit, alkotóelemeit ismernie kell. A szerepkörök változásának, kibővülésének köszönhetően a korábban élesebb szakmai határok fokozatosan elmosódnak.

A modern mérnök számára ezért komoly lehetőséget jelent, ha saját szakterületi tudását az adattudomány eszköztárával egészíti ki. Így nemcsak az adattudósokkal és szoftverfejlesztőkkel való

kommunikáció válik egyszerűbbé, hanem lehetővé válik a közös fejlesztői munka, sőt bizonyos esetekben saját adatvezérelt megoldások létrehozása és integrálása is.

Az adatok és a mérnöki valóság kapcsolata

Tegyük fel, hogy egy mérnök olyan gyártási, folyamattechnikai vagy tervezési problémával találkozik, amelyet hagyományos számításokkal vagy fizikai modellezéssel már nem, vagy csak nagyon nehezen tud megoldani.

Ilyenkor érdemes lehet a statisztikai elemzés és az adatvezérelt modellezés eszköztárához fordulni.

A modern ipari folyamatok jelentős része már nagymértékben digitalizált. Szenzorok, gépek, mérőállomások, vezérlők, adatbázisok és informatikai rendszerek folyamatosan adatokat állítanak elő. Első pillantásra tehát úgy tűnhet, hogy minden rendelkezésünkre áll egy jó modell elkészítéséhez.

A valóságban azonban az, hogy sok adatunk van, még nem jelenti azt, hogy megfelelő információnk is van.

Az egyik legfontosabb feladat ezért a vizsgált folyamat vagy termék megfelelő digitális leképezésének kialakítása.

Ehhez pontosan értenünk kell az adatpontok keletkezésének körülményeit. Tudnunk kell például, hogy egy szenzor mit mér, milyen fizikai mennyiséget reprezentál az adott változó, mikor történik a mérés, milyen mintavételi frekvenciával, milyen pontossággal, és milyen körülmények között.

Ismernünk kell a manuális beavatkozásokat is. Tudnunk kell, mikor avatkozik be egy operátor vagy mérnök a folyamatba, milyen események hatására változhatnak a gyártási/tesztelési paraméterek, hogyan kerülnek az adatok eltárolásra, milyen rendszereken haladnak keresztül, és milyen módosításokon eshetnek át, mire végül az elemzőhöz kerülnek.

Ez azért lényeges, mert az adat önmagában nem azonos a valósággal. Az adat a valós folyamat egy mérési és informatikai rendszeren keresztül létrejövő reprezentációja.

Milyen adatokkal dolgozunk

A mérnöki adattudományban egyetlen termékről vagy folyamatról is sokféle információ állhat rendelkezésre. Ezek különböző forrásokból származhatnak, és a működés más-más részletét írhatják le.

Egy gyártott termékhez például a következő adatok kapcsolódhatnak, zárójelben néhány példával:

- a felhasznált nyersanyagok és beszállítók adatait (anyagkódok, megnevezések),
- a gyártási paramétereket (préselési erő, hőkezelési idő/hőmérséklet),
- a gépek és berendezések állapotát (Rezgés, hőmérséklet, olajnyomás),
- idősoros szenzorméréseket (hőmérséklet, nyomás),
- környezeti adatokat (hőmérséklet, páratartalom),
- minőségellenőrzési eredményeket (geometriai, felületminőségi),
- Képeket (színes kép, szürkeárnyaltos kép, hőkamerás kép, vibrációs spektrum),
- operátori vagy karbantartási információkat (szerszámcseré, leállás, tisztítás),
- valamint a késztermék végellenőrzésének eredményeit (geometriai, funkcionális, vibrációs, erő).

Ezek az adatok a megjelenési formájukban is különböznek. Találkozhatunk pontszerű mérésekkel, idősoros szenzorjelekkel, képekkel, szöveges leírásokkal, eseménynaplókkal vagy összetett mérési görbékkel. Egy pontszerű mérési eredmény egy adott állapotot jellemezhet, míg egy idősor annak

időbeli változásáról is információt hordoz. Idősoros mérés viszont nem csak szenzorból származhat, hanem például az egyes késztermékek adott végellenőrzési értékeinek időbeli sorba rendezéséből is.

Az elemzés szempontjából az a kérdés, hogy ezek közül mely adatok segíthetnek a megoldandó feladatban. Illetve létrehozható-e a különböző adatok kombinációjából új információ. Az adatok kiválasztását, összekapcsolását és modellezésre alkalmas jellemzőkké alakítását a projekt lépéseinél részletesen tárgyaljuk.

Az adattudomány célja: értékteremtés

Az alkalmazott adattudomány elsődleges célja, hogy valós értéket teremtsen egy vállalat, kutatócsoport vagy egy közösség számára az adatok statisztikai elemzésével, jövőbeli események előrejelzésével, illetve összefüggések feltárásával. A matematikailag elegáns algoritmusok és az elméleti módszerek ehhez fontos eszközöket adnak, de egy projekt/feladat sikerét végső soron az határozza meg, hogy milyen szinten sikerült a problémát megoldanunk, és ezzel milyen eredményt értünk el, hasznunkra válik-e a végeredmény, illetve a feladat megoldása során nyert ismeretek, hogyan változtatják a szemléletünket a probléma megoldásával kapcsolatban.

Az elemzésekre épülő javaslatok és megoldások javíthatják a folyamatokat, megkönnyíthetik a munkát, csökkenthetik a költségeket vagy a selejt mennyiségét. Segíthetnek korábban megoldhatatlannak hitt feladatok kezelésében, és akár a szakterületi tudás határait is kitolhatják, oly módon, hogy az eddig nem ismert összefüggésekre fényt derít.

Egy modell vagy elemzés értékét ezért önmagában nem mutatja meg sem a bonyolultsága, sem az előrejelzési pontossága. Azt is meg kell vizsgálnunk, hogy a segítségével jobb döntéseket tudunk-e hozni, és ezek a döntések a gyakorlatban is javulást eredményeznek-e.

Ehhez minden adattudományi projektben az alapoktól kell építkeznünk. A munka gondolatmenete a következő lépésekben foglalható össze.

1. Értsük meg a terméket vagy a folyamatot.
2. Értsük meg az adatpontok keletkezésének körülményeit - illetve a rendelkezésünkre álló adatok jellemzőit.
3. Tegyük fel kérdéseket, és állítsunk fel hipotéziseket.
4. Építsük fel a termék vagy a folyamat digitális történetét, és alakítsuk ki a leíró jellemzőket.
5. Vegyünk fel különböző modellezési "szemüvegeket".
6. Értsük meg az eredményeket: értékeljük a modellek teljesítményét, és végezzünk hibaelemzést.
7. Fordítsuk le az eredményeket a szakterület nyelvére.
8. Értékeljük a szakértőkkel közösen, hogy az eredmények hogyan járulhatnak hozzá a fejlődéshez.
9. Ha szükséges, végezzünk megtervezett, irányított fizikai kísérleteket.
10. Építsünk az eredmények használatát támogató riportoló rendszert vagy alkalmazást.
11. Teszteljük a teljes megoldást.
12. Illesszük be a megoldást a mérnöki munkafolyamatokba.
13. Kövessük nyomon a működését és a hatását.
14. A tapasztalatok alapján fejlesszük tovább.

Ezek a lépések nem merev, egyszer végigjárható sorrendet alkotnak. Egy elemzés új kérdéseket vethet fel, egy teszt hiányzó adatokra világíthat rá, az üzemi tapasztalatok pedig a probléma újrafogalmazását is szükségessé tehetik. Az adattudományi munka ezért folyamatos visszacsatolásra épül, amelyben a folyamat megértése, az adatok vizsgálata és a gyakorlati kipróbálás egymást alakítja. A résztvevők - adatos szakemberek és szakterületi mérnökök - szoros együttműködése (a problémák és elemzési eredmények mély megértéséhez) a sikeresség fontos feltétele.

Az adattudományi projekt lépései a gyakorlatban

A következőkben egy fiktív fémalkatrész gyártásának példáján követjük végig a fenti gondolatmenet tizennégy lépését, és próbáljuk összerakni, milyen módon közelíthetünk meg egy feladatot egy adattudós gondolatmenete alapján. A példa a folyamat megértésétől a kísérletek tervezésén át az üzemi nyomon követésig és továbbfejlesztésig mutatja be az elemzői munkát. Az itt megfogalmazott hipotézisek vizsgálandó feltevések, nem igazolt üzemi eredmények.

Értsük meg a terméket vagy a folyamatot

A szakértő mérnökök gyakran azért fordulnak adattudósokhoz, mert olyan összefüggéseket vagy megoldásokat keresnek, amelyeket a rendelkezésre álló analitikus módszerekkel nehéz vagy akár lehetetlen feltárni. Az együttműködés egyik első feladata ezért a probléma szempontjából lényeges szakterületi tudás megosztása.

Az adattudósnak elegendő ismeretet kell szereznie ahhoz, hogy értelmezni tudja a feladatot, felismerje a fontos körülményeket, és megfelelő elemzési stratégiát alakítson ki. Ez rendszerint több, elmélyült beszélgetést igényel, amely során a szakértő megfelelő szakterületi tudással vértézi fel az adattudóst. A mérnök ismerteti a folyamatot, az elemző visszakérdez -ezzel gyakran olyan pontokra is rávilágítva, amelyet addig talán figyelmen kívül hagytak - és közösen tisztázzák azokat a részleteket is, amelyek a szakterület művelői számára magától értetődők.

Tegyük fel, hogy egy fémalkatrészeket gyártó csoport bizonyos időszakokban a nem megfelelő termékek arányának növekedését tapasztalja. Minden termék végellenőrzésen esik át: egy meghatározott vizsgálati protokoll alapján ellenőrzik az előírt jellemzőit, majd OK vagy NOK (not OK) minősítést kap. Az OK azt jelenti, hogy a termék teljesítette az adott vizsgálat követelményeit. A NOK minősítésű termék javítható lehet; ha javítása nem lehetséges vagy nem indokolt, selejtté válik.

A szakértő ebben a példában három kritikus szakaszt jelöl meg a termék megfelelősége szempontjából: a melegalakításhoz kapcsolódó préselési, edzési és hűtési szakaszt, a temperálást, valamint a végellenőrzés előtti kalibrálást. Ezek adják a vizsgálat kiindulópontját, de később más folyamatlépések jelentősége is felmerülhet.

A préselési, edzési és hűtési szakaszban szenzorok rögzítik például a prés nyomását, a szerszám helyzetét és a hűtőolaj térfogatáramát. A mérnök elmagyarázza, mi a szerepük a folyamatban, milyen működést tekint megfelelőnek, és milyen eltérésekkel találkozott korábban.

A temperálásnál egy energijellegű folyamatjellemző áll az elemzés középpontjában. Már itt tisztázni kell, pontosan mit jelent a tárolt érték: például egy ciklus energiafelhasználását, egy időszak átlagteljesítményét vagy más, a vezérlő által számított mutatót. Ezek különböző fizikai mennyiségek, ezért a változó neve mellett a mértékegységét és a számítási módját is ismernünk kell.

A végellenőrzés megértéséhez a mérnök ismerteti a vizsgált termékjellemzőket, a számítás módját, az elfogadási határértékeket és azok esetleges változásait. Ha egy követelmény időközben módosult,

ugyanaz a mérési eredmény eltérő időpontokban más minősítést is kaphatott. Ezt az elemzésben figyelembe kell vennünk.

Miután megismertük a termék és a folyamat feltételezhetően legfontosabb részeit, megkezdhetjük az adatok első áttekintését. A folyamat megértése az elemzési munka közben is folytatódik: egy szokatlan mérési görbe vagy hiányzó adat újabb kérdéseket vethet fel.

Értsük meg az adatpontok keletkezésének körülményeit

Az adatok értelmezéséhez követnünk kell a keletkezésük útját. Milyen szenzorból származnak? Mit mér a szenzor, és milyen mértékegységben? Mikor és milyen mintavételi gyakorisággal történik a mérés? A rögzített jel mely időszakasza tartozik ténylegesen a vizsgált művelethez, és milyen értéktartományban tekinthető megbízhatónak, illetve mi okozhat hiányos/hibás adatot?

Ugyanilyen fontos, hogy mikor és hogyan avatkozhat be az operátor, hol keletkezik az időbélyeg, illetve folyamatosan vagy adatblokkokban érkeznek-e a mérések. Az adat beérkezési ideje nem feltétlenül egyezik meg a mérés időpontjával. A folyamatos mintavételezésből származó értékeket is továbbbíthatja a rendszer ciklusonként, egyetlen adatcsomagban.

A gyártási példában többféle adat fordul elő. A prés működése közben a szerszám helyzetét, az olaj térfogatáramát és a nyomást szenzorok rögzítik, így egyetlen termék gyártásához is teljes idősor tartozik folyamat paraméterről. Más jellemzőket csak egy-egy meghatározott pillanatban mérnek meg, például a termék hőmérsékletét egy állomás elhagyásakor. Emellett lehetnek ciklusonként összesített mutatók és a végellenőrzés során kapott mérési értékek, valamint minősítések.

Mielőtt összekapcsolnánk a méréseket a különböző gyártási fázisokból, azt is ellenőrizzük, hogyan rendelhetők termékhez, gyártási ciklushoz vagy tételhez. Ha egy alkatrész ismételten áthalad egy mérési állomáson, például újraellenőrzés, vagy javítás és újraellenőrzés miatt, akkor a termékazonosító mellett a művelet vagy a mérési kísérlet azonosítójára is szükség lehet, illetve az időbélyegek segíthetnek sorba rendezni a műveleteket. A meglevő adatokból a lehető legtöbbet próbáljuk meg felhasználni.

Tegyünk fel kérdéseket és állítsunk fel hipotéziseket

Amikor már van képünk a termékről, a folyamatról és az elérhető adatokról, az eredeti problémát konkrétabb kérdésekre bonthatjuk. Ehhez tisztáznunk kell a megrendelő elsődleges célját is: a nemmegfelelőség okait szeretné feltárni, előre szeretné jelezni a hibás termékeket, vagy egy beavatkozási döntéshez vár támogatást?

A feltáró kérdések a folyamat megértését és az adatigény pontosítását segítik. Például:

- Hány beszállítótól érkezik az alapanyag, és visszakövethető-e, melyik termék melyik tételből készült?
- Mi okozhat leállást a gyártósoron, és mely állomások között várakozhat a termék?
- Van-e olyan "vakfolt" a folyamatban, amelyről jelenleg nincs mérésünk, pedig segíthet megérteni a termék minőségében észlelt változásokat és szükségünk lehet ahhoz kapcsolódó információra?

A hipotézis már egy ellenőrizhető feltevést fogalmaz meg. A példánkban ilyen lehet:

- Az eltérő beszállítóktól származó alapanyagokból készített termékek végellenőrzési eredményei vagy NOK arányai különböznek.
- A két gyártósoron a temperálás azonos módon számított energiajellegű mutatójának átlaga eltér, összehasonlítható termékváltozatok és receptek mellett.

- A préselés és a következő hőkezelési lépés között hosszabb időt eltöltő termékeknel gyakoribb a NOK minősítés esetlegesen nem optimális hőállapotok kialakulása miatt.

Ezeket a feltevéseket a vizsgálat előtt tovább pontosítjuk: melyik termékváltozatra, időszakra és minőségi jellemzőre vonatkoznak, illetve mit tekintünk érdemi eltérésnek? A "hosszabb várakozás" például jelenthet egy technológiai indokolt küszöb átlépését, de vizsgálhatjuk az eltelt idő teljes tartományát is.

Egy megfigyelt különbség többféle magyarázattal összeegyeztethető. Ha az egyik beszállító alapanyagát főként az egyik soron használják, a beszállító és a gyártósor szerepét nehéz lehet elkülöníteni. A megfigyelésekből kiinduló magyarázatot ezért külön kell választanunk attól, hogy egy beavatkozás ténylegesen megváltoztatná-e az eredményt. Erre a modellezési szemléleteknél még visszatérünk.

Építsük fel a digitális történetet és a leíró jellemzőket

A termék digitális története az adott termékhez rendelhető gyártási események és mérések összekapcsolt leírása. A cél az, hogy követhetővé váljon, mi történt a termékkel a folyamat során, és milyen körülmények között alakult ki a végellenőrzésen megfigyelt állapota.

A történet felépítéséhez ellenőriznünk kell, hogy egy mérés valóban a megfelelő termékhez és gyártási ciklushoz tartozik-e, illetve a rögzített események időrendje helyes-e. Egy hibás összerendelés vagy félreértelmezett időbélyeg a későbbi következtetéseket is torzíthatja.

A sok adat akkor válik elemezhetővé, ha a kérdésünkhöz és a választott módszerhez megfelelő formába rendezzük. A példánk egyik lehetséges kiindulópontja egy olyan táblázat, amelyben minden sor egy terméket, minden oszlop pedig annak egy jellemzőjét tartalmazza. A sorhoz hozzákapcsoljuk a megfelelő gyártási eseményeket és a végellenőrzés eredményét is.

Ehhez a különböző részletességű adatokat közös szintre kell hoznunk. Egy pontszerű hőmérsékletmérés közvetlenül lehet a termék egyik jellemzője. Egy teljes préselési idősből viszont több leíró értéket is kiszámíthatunk. A nyomásjelből például maximumot, átlagot, szórást vagy egy kijelölt nyomásszint fölött eltöltött időt képezhetünk. A szerszám mozgásából a véghelyzet vagy egy mozgásszakasz időtartama lehet érdekes.

Nem kell minden idősből egyetlen számot készítenünk. Több jellemző együttesen írhatja le a görbe különböző tulajdonságait. Itt a táblázatos modellezéshez alakítjuk át az adatokat; léteznek olyan megközelítések is, amelyek a jelsorozatot közvetlenebb formában dolgozzák fel.

A jellemzőkinyerés részben automatizálható. A tsfresh Python-csomag például számos statisztikai és jelleíró tulajdonságot számít ki idősorokból, amelyek később regressziós, osztályozási vagy csoportosítási feladatokban használhatók. Az automatikus számításnál is ellenőrizni kell a bemenetet és a módszerek feltételezéseit, például azt, hogy szükség van-e egyenletes mintavételezésre. [1]

Mérnöki iránymutatással célzott jellemzőket is kialakíthatunk. Ha például az olaj térfogatáramának egy kijelölt jelszakasza csillapodó lengést mutat, megvizsgálhatjuk, leírható-e csillapított rezgést közelítő függvényvel. A becsült amplitúdó, frekvencia és csillapodási paraméter külön jellemzőként szerepelhet. A rosszul illeszkedő görbék paramétereit azonban nem célszerű ugyanolyan megbízható értékeként kezelni. Maga az illeszkedés sem bizonyítja, hogy a jel mögött pontosan a feltételezett fizikai mechanizmus áll.

Az időbélyegegk további jellemzők kialakítását teszik lehetővé. Kiszámíthatjuk például a préselés befejezése és a temperálás kezdete között eltelt időt:

Átadási idő = a temperálás kezdő időpontja – a préselés befejezési időpontja.

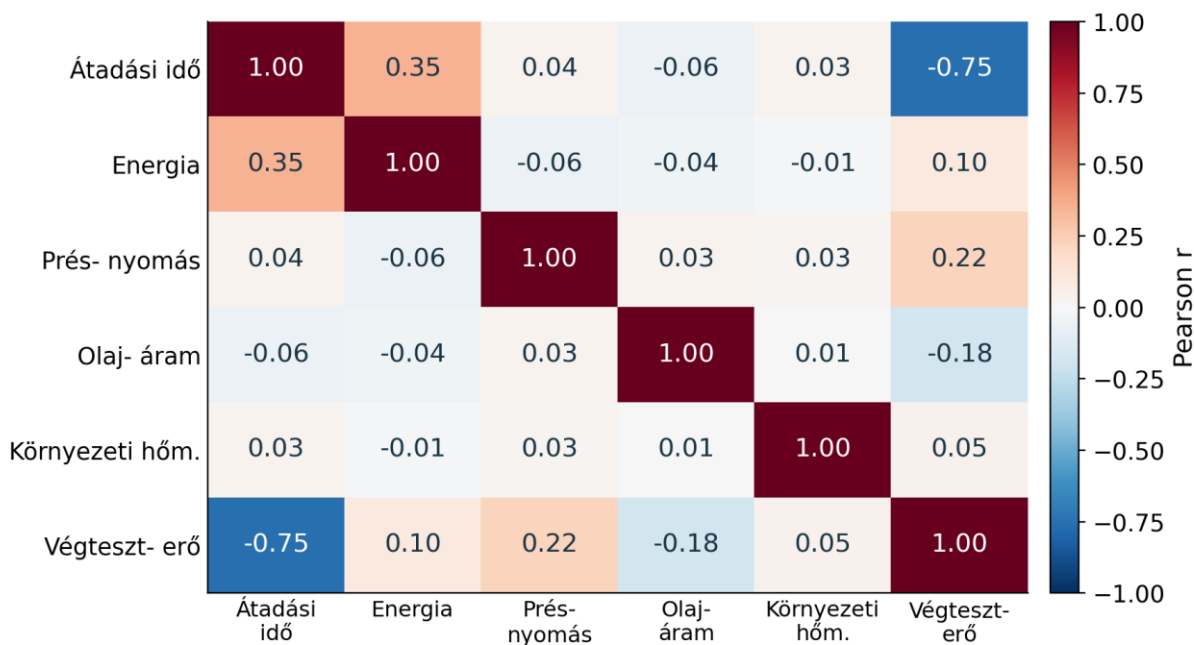
A képlet egyszerű, de csak akkor kapunk értelmezhető eredményt, ha ugyanahhoz a termékhez tartozó, megfelelő eseményeket kapcsolunk össze, és a két rendszer időbélyegei összehasonlíthatók. Ez az idő a szállítást és a várakozást is tartalmazhatja; önmagában nem azonos a sor leállási idejével.

Az időbélyegek alapján csoportokat is képezhetünk: vizsgálhatunk műszakokat, napokat vagy termékváltozatokként elkülönített időszakokat. Így megfigyelhetjük, mikor nő a NOK arány, vagy mikor változik meg egy mérési jellemző eloszlása. A NOK darabszáma mellé a vizsgált termékek számát is oda kell tennünk, mert több gyártott darab mellett változatlan hibaarány esetén is több hibás darabot találhatunk.

A meglévő adatokból új, az elemzés számára használható jellemzők kialakítását jellemzőképzésnek, angolul feature engineeringnek nevezzük. Ez magában foglalhat statisztikai összesítést, átalakítást, változók kombinálását és szakterületi összefüggések felhasználását. A témát részletesen tárgyalja Max Kuhn és Kjell Johnson Feature Engineering and Selection című könyve. [2] A feature-ök, illetve a célrétékek között számolt korrelációt, korrelációs mátrix hőterképével szemléltethetjük hatékonyan. Az alábbi ábrán (1. ábra Korrelációs mátrix) lineáris Pearson korreláció van szemléltetve, ahol a meleg színek a pozitív korrelációt, a hideg színek pedig a negatív korrelációt mutatják. Pozitív korreláció esetében, a sorbarendezt értékek monoton növekvő trendet mutatnak, nagysága pedig az értékekre illesztett egyenes körüli szórásának alapján értelmezhető. Minél jobban illeszkedik az egyenes az értékekre- annál nagyobb a korreláció értéke. Negatív korreláció esetében viszont az egyik érték csökkenésénél a másik növekszik. A folyamat, termékjellemzők egymásra gyakorolt hatásáról lehet következtetéseket levonni a korrelációk elemzésével.

1. Korrelációs mátrix

Szintetikus gyártási adatok · 1100 tanítópélda · Pearson-féle r



A szín a lineáris kapcsolat irányát és erősségét mutatja.

A korreláció önmagában nem bizonyít ok-okozati kapcsolatot.

Előrejelző feladatnál már az adatok összeállításakor meg kell határoznunk, a folyamat mely pontján szeretnénk becslést adni. A modell bemenetei közé csak olyan információ kerülhet, amely a későbbi használat során ebben az időpontban ténylegesen elérhető lesz. Ha például a hőkezelés előtt szeretnénk előre jelezni a termék végső minőségét, akkor az adott termék hőkezelése során mért adatokat még nem használhatjuk fel. A tanítóadatok összeállításakor is ezt a felhasználási helyzetet kell követnünk: minden terméknel az előrejelzés tervezett időpontjáig elérhető adatok adják a bemeneteket, a később mért végső minőség pedig a megtanulandó célértéket. Ha a célértékre valamilyen formában utaló eredmény bekerül a tanító táblába akkor információszivárgás keletkezik, amely megtévesztően jó kiértékelési eredményt adhat. [11]

Vegyünk fel különböző modellezési szemüvegeket

A modellezési megközelítést az határozza meg, milyen kérdésre keressük a választ. Más feladat két csoport jellemzőinek összehasonlítása, egy mérési eredmény előrejelzése, a hasonló termékek csoportosítása vagy egy beavatkozás hatásának vizsgálata.

Ezeknek a szemléleteknek az áttekintéséhez Christoph Molnar Modeling Mindsets című könyve ad támpontot. A kötet többek között a gyakorisági és Bayes-i következtetés, az oksági elemzés, valamint a felügyelt és felügyelet nélküli tanulás nézőpontjait mutatja be. [3]

Az alábbi megközelítések nem mind ugyanazon szempont szerint felállított, egymást kizáró kategóriák. Egy regressziós modell például lehet felügyelt tanulási feladat része, és becsülhetjük a paramétereit Bayes-i módon is.

Hipotézisvizsgálat

Hipotézisvizsgálattal azt értékeljük, hogy az adataink mennyire egyeztethetők össze egy meghatározott nullhipotézissel, a választott statisztikai eljárás feltételezései mellett. A két gyártósor összehasonlításakor a nullhipotézis például az lehet, hogy az energijellegű mutató sokasági átlaga azonos. Megfelelő körülmények között ezt kétféleképpen t-próbával vizsgálhatjuk. [4]

Ha nem mutatunk ki szignifikáns eltérést, az még nem bizonyítja a két csoport azonosságát. Az eredmény függhet a minta méretétől, a változékonyságtól és a vizsgálat feltételeinek teljesülésétől. A statisztikai döntés mellett a becsült különbség nagyságát és bizonytalanságát is értékeljük. [4]

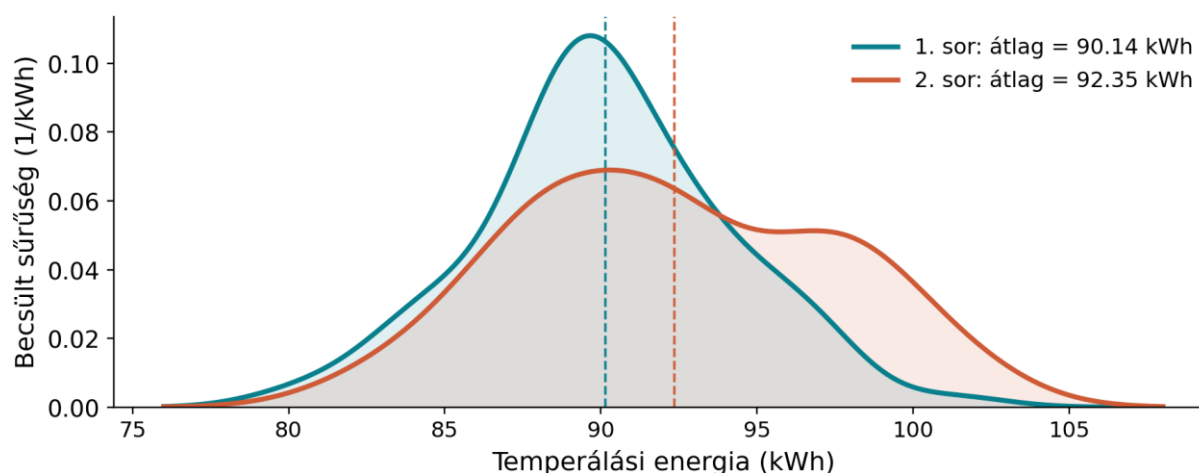
A gyártási példában az is kérdés, hogy a talált különbség mérnöki szempontból számít-e. Egy kis átlageltérés önmagában nem feltétlenül indokol technológiai beavatkozást. Ezzel szemben egy bizonytalan becslés is indokolhat további mérést, ha a lehetséges eltérés következménye jelentős.

Hipotézis vizsgálat elvégzése esetében hatékonyan lehet használni az adatok eloszlásának megjelenítésére a valószínűségi sűrűség függvény (probability density function) számolt görbéket (2. ábra Két gyártósor összehasonlítása). A t-próba (t-test) során számolt p-érték alapján dönthetjük el, hogy a két adathalmaz eloszlása statisztikailag szignifikáns-e vagy sem. Általánosan elfogadott szokásnak a $p=0.05$ értéket használják, amely érték alatt a két eloszlás átlaga statisztikailag eltérő, felette viszont nem.

2. Két gyártósor összehasonlítása

Szintetikus, független megfigyelések · 80 termék / sor · kétoldali Welch-féle t-próba

$$H_0: \mu_1 = \mu_2 \quad t = 3.08 \quad p = 0.0025$$



Átlagkülönbség (2. sor – 1. sor): 2.20 kWh; 95%-os CI: [0.79; 3.62].

A KDE simított eloszláskép. A t-próbát az eredeti mintákon számoltuk.

Felügyelt tanulás

Felügyelt tanulásnál (supervised learning) olyan példákban tanulunk, amelyekhez ismerjük a bemeneti jellemzőket és a becsülni kívánt kimenetet. Ezt a kimenetet célváltozónak (target variable), statisztikai szóhasználatnál függő változónak (dependent variable) nevezzük. A modell feladata, hogy a tanult összefüggést új esetekre is alkalmazhatóvá tegye. [5]

Regresszióval egy számszerű kimenetet becsülhetünk (3. ábra Regressziós előrejelzés és hibaelemzés). A példánkban a végellenőrzés egyik mért termékjellemzőjét szeretnénk előre jelezni a korábbi gyártási lépések adataiból. Osztályozással, más néven klasszifikációval kategóriát becsülünk: például azt, hogy egy termék OK vagy NOK minősítést kap-e. Egy megfelelő modell a NOK esemény becsült valószínűségét is megadhatja, amelyből döntési szabállyal képezhetünk címkét. [5]

A bemenetek között szerepelhetnek a préselési görbe leíró jellemzői, az átadási idő és a hőkezelés mért adatai. A bemenetek köre attól is függ, hogy a folyamat melyik pontján akarunk előrejelzést készíteni.

A modellek az adataink megértését is segíthetik: megvizsgálhatjuk, milyen bemeneti mintázatokhoz társítanak kedvező vagy kedvezőtlen kimenetet. A prediktív összefüggés azonban önmagában nem bizonyítja, hogy az adott változó módosítása megváltoztatná a termék minőségét. Az előrejelzés és az oksági magyarázat eltérő cél, amely eltérő modellezési döntéseket is igényelhet. [7]

Felügyelet nélküli tanulás

Felügyelet nélküli tanulásnál (unsupervised learning) nem adunk meg olyan célcímkét, amelynek előrejelzését a modellnek meg kellene tanulnia. Az adatok szerkezetét vizsgáljuk. Ennek egyik lehetősége a klaszterezés, vagyis hasonló megfigyelések csoportokba rendezése. A csoportok a választott jellemzőktől és módszertől is függenek. [6]

A gyártási példában a préselési görbékből számított jellemzők alapján képezhetünk termékcsoportokat, majd utólag megvizsgálhatjuk, miben különböznek. Kapcsolódnak-e eltérő gyártósorokhoz,

alapanyagokhoz vagy üzemállapotokhoz? Eltér-e a NOK arányuk? A csoportosításba ekkor nem adjuk be az OK/NOK címkét, de a létrejött csoportok értelmezéséhez felhasználhatjuk.

Bizonyos módszerek, például a k-közép (k-means), új termékeket is hozzá tudnak rendelni a már kialakított csoportokhoz. Ez azonban nem minden klaszterező eljárás beépített képessége; a későbbi alkalmazás módját a módszer kiválasztásakor kell tisztáznunk. [6]

Bayes-i modellezés

A Bayes-i megközelítésben az ismeretlen modellparaméterekre vonatkozó kezdeti feltevéseinket és bizonytalanságunkat egy előzetes, úgynevezett prior eloszlással írjuk le. A megfigyelési modell azt fejezi ki, hogyan keletkezhetnek az adatok a paraméterek adott értékei mellett. Az adatok figyelembevételével frissített paramétereloszlást posterior eloszlásnak nevezzük. [8]

A példánkban Bayes-i regresszióval is vizsgálhatjuk, hogyan függ össze a végellenőrzés mért értéke a préselési jellemzőkkel és az átadási idővel. A paraméterek becslése mellett azok bizonytalanságát is bemutatjuk. A priorok megadásánál szakterületi ismeretekre támaszkodhatunk, de a feltevések hatását és a modell illeszkedését ellenőriznünk kell. Ez a megközelítés sem teszi automatikusan oksággá a feltárt kapcsolatokat. [7, 8]

Több nézőpont együttes használata

Érdemes több megközelítést mérlegelni, és a feladathoz illő módszereket összehasonlítani. A megszokott módszer jó kiindulópont lehet, de nem biztos, hogy minden kérdésre az adja a leghasznosabb választ. A különböző szemléletek erősségeinek és korlátainak felismerése a modellezői eszköztár fontos része. [3]

Ha több módszer hasonló mintázatot mutat, az indokolhatja annak további vizsgálatát. Ugyanakkor közös adatkezelési hiba, mintavételi torzítás vagy egy figyelmen kívül hagyott folyamatjellemző több elemzés eredményét is befolyásolhatja. Az egybehangzó eredmény tehát nem önálló bizonyíték a fizikai magyarázat helyességére. Az oksági állításhoz megfelelő vizsgálati terv és feltevések szükségesek. [7]

A példánkban ezért egy következő lépés lehet az átadási idő megtervezett változtatása összehasonlítható gyártási körülmények között. Ez már irányított üzemi kísérlet lenne, nem a historikus adatok újabb algoritmussal történő elemzése. A kísérlet célját, változtatott tényezőit és elrendezését előre kell meghatározni. [12]

Értsük meg az eredményeket

George E. P. Box ismert megfogalmazása szerint: "Minden modell hibás, de némelyik hasznos". Ez az "All models are wrong but some are useful" mondat saját magyar fordítása. [9, 202. o.]

Egy modell a vizsgált folyamat közelítő leírása. A historikus adatokból felépített statisztikai modell használhatóságát ezért az adott célhoz mérjük: milyen pontossággal, milyen körülmények között és mekkora bizonytalansággal ad választ? A jó illeszkedés mellett az eltérések vizsgálata is a modellépítés része. [9]

Tanítóadatok validációs adatok és tesztadatok

Felügyelt tanulásnál három szerepet érdemes elkülöníteni. A tanító adathalmaz (training set) alapján illesztjük a modellt. A validációs adathalmaz (validation set) a fejlesztés közbeni választásokat, például a modellbeállítások összehasonlítását segíti. A teszt adathalmaz (test set, evaluation set) a kiválasztott megoldás végső, elkülönített kiértékelésére szolgál. [10]

Nem szükséges az adatokat egyszer és végleg három részre osztanunk. A tanításhoz és a modellválasztáshoz keresztvalidációt is használhatunk: ilyenkor több körben dolgozunk, és körönként más adatrészen tanítjuk, illetve ellenőrizzük a modellt. A végső tesztalmazt azonban ebből a folyamatból kihagyjuk, és csak a kiválasztott modell teljesítményének értékelésére használjuk. Tegyük fel, hogy elkészítünk egy modellt, majd megnézzük, hogyan teljesít a tesztalmazon. Az eredmény alapján megváltoztatjuk a beállításait, új bemeneti jellemzőket választunk, vagy másik algoritmust próbálunk ki. Ha ezután ugyanazon a tesztalmazon ellenőrizzük, javult-e a teljesítmény, akkor már a tesztadatokhoz is igazítjuk a megoldásunkat - még akkor is, ha közvetlenül nem tanítjuk rajtuk a modellt. Sok ilyen próbálkozás után olyan változatot választhatunk, amely részben az adott tesztalmaz sajátosságai miatt teljesít jól, és új adatokon már gyengébb eredményt ad. Ezért a fejlesztési döntéseket a validációs eredményekre alapozzuk, a végső tesztalmazt (evaluation set) pedig a már kiválasztott megoldás ellenőrzésére tartjuk fenn. [10]

Gyártási adatoknál az időbeli és csoportkapcsolatokra is figyelni kell. Ha jövőbeli gyártást akarunk előre jelezni, a korábbi időszakokból történő tanítás és a későbbiekben végzett teszt közelebb állhat a tényleges alkalmazáshoz. Ugyanazon termék ismételt méréseit pedig együtt kell kezelni a felosztásban, ha új termékeken várható teljesítményt szeretnénk becsülni. [10]

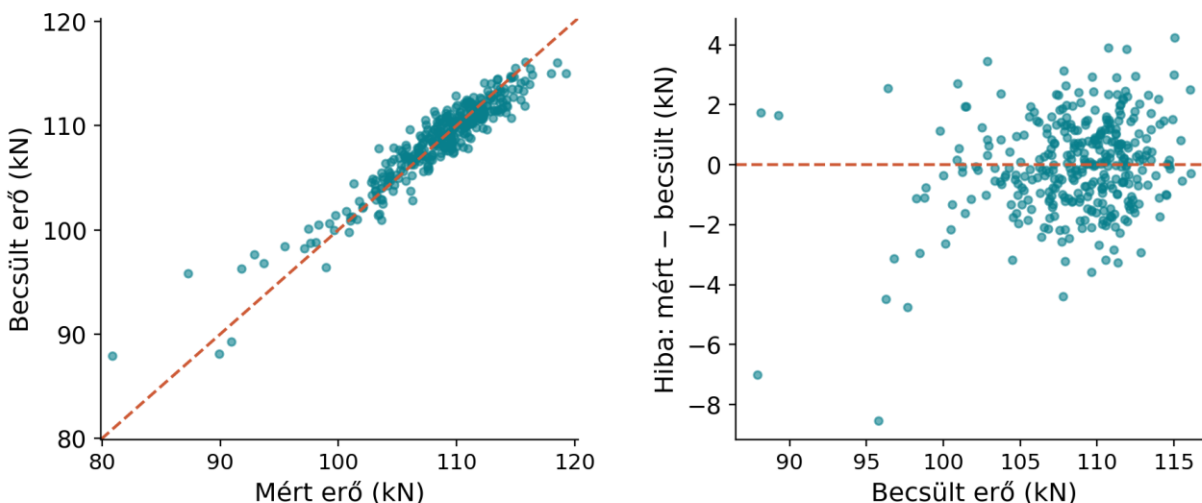
Egyes modellek alkalmazása előtt átalakítjuk a bemeneti változók számszerű értékeit. Például standardizáláskor kivonjuk egy változó átlagát, majd elosztjuk az értékeket a szórásával. Az ehhez szükséges átlagot és szórást a tanítóadatokból számoljuk ki, és a többi adathalmazra is ugyanezt az átalakítást alkalmazzuk. Keresztvalidáció esetén ezt minden körben az adott tanítórész alapján végezzük el. Hasonlóképpen, ha az adatok alapján választjuk ki a modell bemeneti jellemzőit, abba sem vonhatjuk be a végső tesztadatokat. [11]

A hibaelemzés kezdete

A kiértékelés során azt is meg kell értenünk, milyen esetekben téved a modell. A gyártási példában külön megnézhetjük az egyes gyártósortokat, termékváltozatokat és időszakokat. Regressziónál a mért és becsült értékek eltéréseit vizsgáljuk. Osztályozásnál megkülönböztetjük a valójában NOK, de OK-nak becsült termékeket azoktól, amelyek megfelelőek, mégis NOK jelzést kapnak. A két tévedés következménye eltérő lehet, ezért a modell gyakorlati hasznát ezek ismeretében kell megítélnünk.

3. Regressziós előrejelzés és hibaelemzés

Szintetikus példa · Random Forest · 350 elkülönített tesztermék · $R^2 = 0.891$



Átlagos abszolút hiba (MAE): 1.20 kN.

Balra a tökéletes becslés egyenese; jobbra a hibák eloszlása a becslések mentén.

A klasszifikációs modellek teljesítményének értékelése

Klasszifikáció esetén a modell által előre jelzett címkéket a tényleges címkékkel hasonlítjuk össze. Először egyértelműen meg kell határoznunk, melyik osztályt tekintjük pozitívnak. A gyártási példánkban legyen ez a NOK, vagyis a nem megfelelő termék. A "pozitív" itt az általunk keresett eseményt jelöli. Az eredményeket tévesztési mátrixban (confusion matrix) (4. ábra Tévesztési mátrix) foglalhatjuk össze. [13]

Tényleges állapot	Modell szerint NOK	Modell szerint OK
NOK	Valódi pozitív (TP) Felismert nemmegfelelőség	Hamis negatív (FN) Fel nem ismert nemmegfelelőség
OK	Hamis pozitív (FP) Téves riasztás	Valódi negatív (TN) Helyesen felismert megfelelő termék

Az esetszámokból több teljesítménymutatót számíthatunk. Ezek különböző kérdésekre válaszolnak, ezért együtt érdemes értelmezni őket. Az alábbi precision, recall és F1 értékek a NOK-osztályra vonatkoznak. [13]

Metrika	Képlet	Mérnöki értelmezés
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$	Az összes termék mekkora részét soroltuk helyesen osztályba?
Recall	$TP / (TP + FN)$	A ténylegesen NOK termékek mekkora részét ismertük fel?
Precision	$TP / (TP + FP)$	A NOK-nak jelzett termékek mekkora része valóban NOK?

Metrika	Képlet	Mérnöki értelmezés
F1-score	$2TP / (2TP + FP + FN)$	A precision és a recall harmonikus középértéke.

A képletekben nem nulla nevezőt feltételezünk. Ha például a modell egyetlen terméket sem jelez NOK-nak, a precision közvetlenül nem számítható; a kiértékelésnél jelezzük az ilyen eseteket. [13]

4. Tévesztési mátrix

Szintetikus példa · 350 tesztermék · NOK = pozitív osztály · küszöb: 0,50

Valós OK	TN 329 db	FP 1 db
Valós NOK	FN 4 db	TP 16 db
	jelzett OK	jelzett NOK

Precision = 94.1% Recall = 80.0% F1 = 0.865
FP: téves NOK-jelzés; FN: fel nem ismert NOK-termék.

Kiegyensúlyozatlan osztályok és a hibák költsége

Gyártási adatoknál gyakori, hogy sokkal több megfelelő termékünk van, mint nem megfelelő. Ilyenkor kiegyensúlyozatlan osztályokról (imbalanced data) beszélünk. A kevés NOK-példa korlátozhatja a hozzájuk tartozó mintázatok megtanulását, de a feladat nehézségét az osztályarány mellett az adatok információtartalma is meghatározza.

Tegyük fel, hogy 1000 termékből 990 OK és 10 NOK. Egy olyan modell, amely minden termékre OK-t jelez, 99%-os accuracy értéket ér el. Eközben egyetlen NOK-terméket sem ismer fel: a NOK-osztály recall értéke nulla. A magas accuracy tehát önmagában félrevezető képet adhat a modell használhatóságáról. [13]

A megfelelő kiértékeléshez azt is tisztáznunk kell, milyen következménye van az egyes tévedéseknek. A hamis pozitív jelzés felesleges ellenőrzést vagy termékviSSzatartást okozhat. Hamis negatív jelzés miatt viszont elmaradhat egy szükséges beavatkozás. A következmény attól is függ, milyen ellenőrzések maradnak a folyamatban.

A metrikát ezért a tervezett felhasználáshoz kell választanunk. Ha számszerűsíthetők a következmények, egyszerű költségfüggvényt is felállíthatunk: $C = c_{FP} \times FP + c_{FN} \times FN$. A c_{FP} és c_{FN} a kétféle tévedés egy esetre jutó költsége. Ez egyszerűsített leírás: az ellenőrzések és beavatkozások további költségeit is figyelembe vehetjük. Az F1-score önmagában nem tartalmazza ezeket a gazdasági következményeket. [14]

A döntési küszöb megválasztása

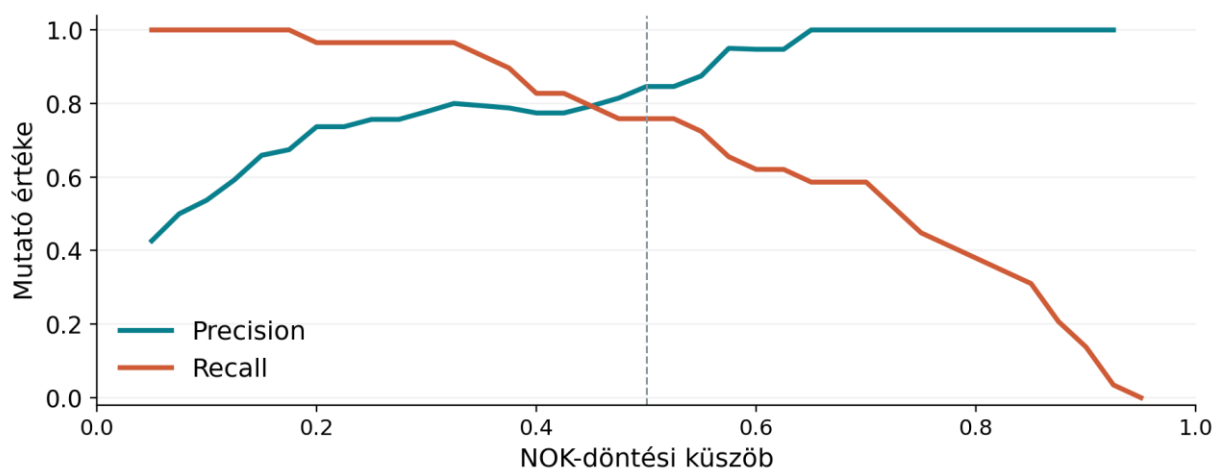
Sok klasszifikációs modell az osztálycímke mellett valószínűségi becslést is ad. Ebből egy döntési küszöb segítségével képezhetünk OK vagy NOK címkét (5. Ábra Döntési küszöb, precision és recall). Tegyük fel, hogy egy termékre a modell 0,3-as NOK-valószínűséget becsül. Ha 0,5-ös küszöböt használunk, OK-nak minősítjük. Ha a küszöböt 0,2-re csökkentjük, ugyanazt a terméket már NOK-nak jelezzük. A modell becslése eközben változatlan; a hozzá kapcsolt döntési szabály módosul. [15]

Alacsonyabb küszöbnél több terméket jelölhetünk NOK-nak. Ezzel több tényleges nemmegfelelőséget ismerhetünk fel, ugyanakkor több téves riasztást is kaphatunk. A küszöböt a kívánt felismerési arány, az ellenőrzési kapacitás és a hibák költsége alapján érdemes megválasztani. [15]

A küszöböt validációs adatokon vagy megfelelő keresztvalidációval állítsuk be. A végső tesztadatokat tartsuk fenn a már kiválasztott modell és döntési szabály független értékelésére. [10, 15]

5. Döntési küszöb, precision és recall

Szintetikus példa · 350 validációs termék · ugyanaz a rögzített klasszifikátor



NOK-jelzés nélkül a precision nem értelmezhető; ott a görbe megszakad.
A választott szabályt ezután külön tesztadatokon értékeljük.

Fordítsuk le az eredményeket a szakterület nyelvére

A szakterületi szakértőknek és döntéshozóknak azt kell megérteniük, hogy a modell teljesítménye milyen gyakorlati eredményt tesz lehetővé. Az F1-score mellett ezért olyan kérdésekre is választ kell adnunk, mint hogy hány nem megfelelő terméket ismerhetünk fel időben, mennyi további ellenőrzésre lesz szükség, és milyen veszteséget előzhetünk meg.

A várható javulást a jelenlegi folyamathoz kell viszonyítanunk. A historikus adatokon elért jó predikciós teljesítmény még nem mutatja meg közvetlenül, mekkora javulást okoz majd a modell használata. Ehhez azt is meg kell határoznunk, milyen intézkedés követi az előrejelzést, és mennyire hatásos ez az intézkedés.

Modelleket a folyamatok jobb megértésére is készítünk. Ebben segíthetnek az értelmezhető gépi tanulás és a magyarázható mesterséges intelligencia (Explainable AI, XAI) módszerei. Használhatunk eleve könnyen áttekinthető modelleket, például egyszerű döntési fákat, vagy összetettebb modellekhez kapcsolódó utólagos magyarázó módszereket. [16]

A globális értelmezés a modell általános viselkedését vizsgálja: mely jellemzőkre támaszkodik, és milyen mintázatokat használ. A lokális értelmezés egy adott termék előrejelzését magyarázza. A magyarázatok a modell ellenőrzését is segítik: megmutathatják, ha az váratlan vagy szakmailag nehezen indokolható összefüggésekre épül. [16]

A modell értelmezése SHAP segítségével

Térjünk vissza a gyártási példánkhoz, azon belül a regressziós feladathoz: most a végtesztelés egyik folytonos mérési értékét szeretnénk előre jelezni. Tegyük fel, hogy a Random Forest regressziós modellünk egy megfelelően elkülönített tesztadathalmazon $R^2 = 0,9$ értéket ér el. Ez azt jelenti, hogy a négyzetes hibák összege az adott tesztadathalmazon a tényleges célértékek átlagával készített állandó becslés hibájának 10%-a. Ez kedvező predikciós eredmény, de nem jelent minden egyes terméknel 90%-os pontosságot. [17]

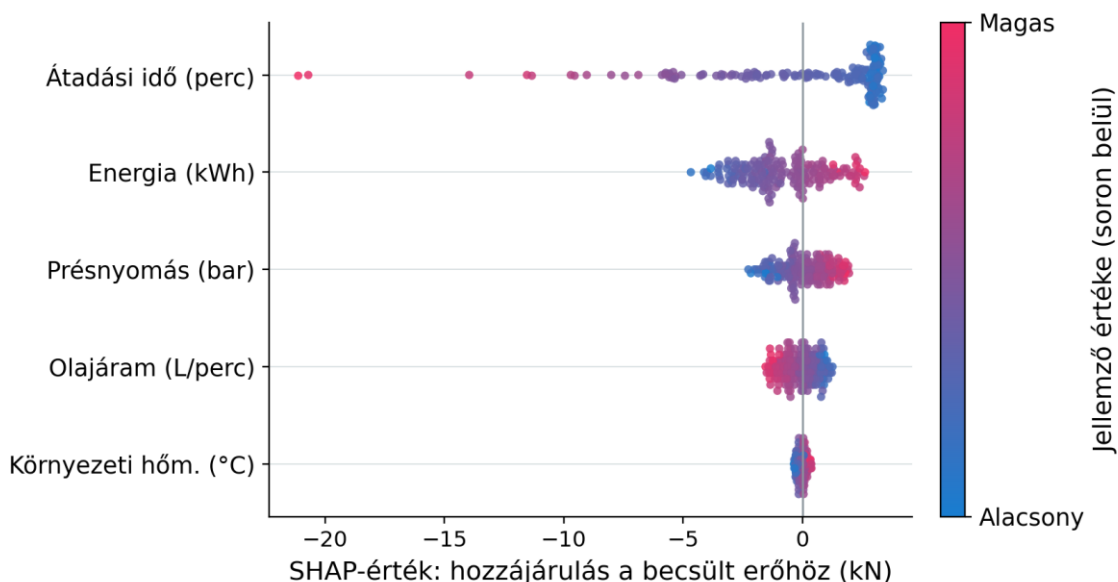
A modell viselkedését például SHAP segítségével vizsgálhatjuk. A módszer a játékelméleti Shapley-értékekre épül: az egyes jellemzőkhöz hozzájárulásokat rendel, amelyek egy referenciaértékhez képest felbontják az adott előrejelzést. [18]

A célváltozó eredeti skáláján magyarázott regressziós kimenetnél a pozitív SHAP-érték növeli, a negatív csökkenti a becsült értéket a referenciához képest. A referencia a választott háttéradatoktól és magyarázati beállításoktól függ; gyakran a modell háttéradatakon vett átlagos kimenetéhez kapcsolódik. [18]

Egy termék előrejelzését külön is elemezhetjük, illetve sok termék SHAP-értékeit közös ábrán vizsgálhatjuk (6. Ábra SHAP summary plot). Így láthatóvá válhat, hogy például a hosszabb várakozási idő milyen irányú hozzájárulással társul a modellben. [18]

6. SHAP summary plot

Szintetikus példa · Random Forest regresszió · 180 teszttermék · 32 háttérminta



Egy pont egy termék. Sorrend: átlagos abszolút SHAP-érték.

Interventional Shapley-felbontás; a modell hozzájárulásai nem oksági hatások.

Az összefüggéstől a mérnöki hipotézisig

Tegyük fel, hogy az elemzésben a préselés és a hőkezelés között eltelt idő fontos jellemzőnek bizonyul. A hosszabb várakozásokhoz olyan SHAP-értékek társulnak, amelyek a becsült végtesztértéket a megengedett tartomány kedvezőtlen szélé felé módosítják. Ezt a szakértő mérnökkel közösen például a következő hipotézissé alakíthatjuk:

Ha a termék túl sokáig vár a hőkezelő állomás előtt, megváltozhat a belépési hőmérséklete. Amennyiben a hőkezelési folyamat ezt nem kompenzálja megfelelően, az eltérő hőállapot hozzájárulhat a végső tulajdonságok kedvezőtlen alakulásához.

A SHAP-elemzés azt mutatja meg, hogyan használja a modell a várakozási időt az előrejelzésben. A lehűlésen alapuló fizikai magyarázat egy további, ellenőrzendő hipotézis. A jó predikciós teljesítmény önmagában nem igazolja ezt az oksági kapcsolatot. [19]

A következő lépésben visszakereshetjük a hosszú várakozással érintett termékeket. Példánkban azt találjuk, hogy több ilyen időszakban egymást követő termékek végtesztértékei is a határértékeken kívülre kerültek. Ellenőriznünk kell azonban, hogy ezekben az időszakokban változott-e más körülmény is, például az alapanyag, a gépbeállítás vagy a gyártott termékváltozat. A közös sorleállítás által érintett termékek hasonló eredményei közös esemény következményei is lehetnek.

A továbblépéshez megvizsgálhatjuk a hőkezelés előtti tényleges hőmérsékletet, összevethetünk hasonló gyártási körülmények között készült termékeket, illetve szükség esetén irányított fizikai kísérletet tervezhetünk. Így a modellezési eredményből konkrét, ellenőrizhető mérnöki vizsgálat lesz. [12]

Értékeljük a szakértőkkel közösen az eredmények gyakorlati hasznát

Az eredmények szakterületi értelmezése után közösen kell megvizsgálnunk, hogy azok milyen fejlődést tehetnek lehetővé. Egy jól teljesítő modell gyakorlati értéke attól függ, hogyan használjuk fel az előrejelzéseit vagy az elemzés során szerzett ismereteket.

A prediktív modellek egyik fontos alkalmazása a várható nemmegfelelőség korai felismerése. Ha már a gyártási folyamat közben jelezni tudjuk, hogy egy termék várhatóan nem lesz megfelelő, a mérnöknek lehetősége nyílna beavatkozni. A folyamat adottságaitól függően módosíthatja a további gyártási paramétereket, kiegészítő ellenőrzést vagy javítást kezdeményezhet, illetve indokolt esetben leállíthatja a termék további feldolgozását.

Ezzel elkerülhetővé válhat a későbbi műveletek egy része, ami időt, energiát és szerszámelettartamot takaríthat meg. Ehhez azonban az előrejelzésnek még a beavatkozás előtt rendelkezésre kell állnia, és valóban befolyásolható folyamatlépéshez kell kapcsolódnia. A téves riasztások miatt végzett felesleges ellenőrzés, javítás vagy selejtezés költségét szintén figyelembe kell vennünk.

Az értékteremtés másik lehetősége az elemzői munka automatizálása. Az ismétlődő adat-előkészítés, tisztítás, ábrakészítés és riportolás automatizálásával időt takaríthatunk meg a felhasználónak. Így több figyelmet fordíthat az eltérések értelmezésére és a szükséges intézkedések megtervezésére.

A modellekből és elemzésekből szerzett ismeretek a folyamat jobb megértését is segíthetik. Új összefüggésekre vagy korábban kevésbé vizsgált körülményekre irányíthatják a szakértők figyelmét. Az ellenőrzött felismerések alapján finomíthatók lehetnek a gyártási beállítások, a folyamatfelügyelet vagy a mérési rendszer. A példánkban ilyen irány lehet a hőkezelés előtti várakozás és belépési hőmérséklet együttes vizsgálata.

A közös értékelés során ezért előre rögzítsük, mit tekintünk javulásnak: kevesebb selejtet vagy utómunkát, kisebb energiafelhasználást, rövidebb átfutási időt, kevesebb kézi elemzői munkát vagy egy

korábban bizonytalan folyamatjelenség jobb megértését. Határozzuk meg a kiindulási állapotot és azt is, hogyan fogjuk a változást mérni. A várható haszon ekkor még becslés; a tényleges eredményt a kipróbálás és a későbbi nyomon követés mutatja meg.

Ha szükséges végezzünk megtervezett irányított fizikai kísérleteket

Az eddigi elemzések megmutathatják, hol hiányosak az ismereteink, milyen körülmények között bizonytalan a modellünk, és mely feltételezett összefüggéseket érdemes célzottan ellenőrizni. Ilyenkor a meglévő adatok további elemzése mellett új mérésekre vagy megtervezett gyártási kísérletekre is szükség lehet.

Bővítsük célzottan a rendelkezésre álló adatokat

Ha a modell bizonyos bemeneti értékeknél vagy azok egyes kombinációinál rendszeresen nagyobbban téved, az jelezheti, hogy az adott tartományban korlátozottan használható. Ennek egyik lehetséges oka, hogy hasonló esetek ritkán szerepeltek a tanítóhalmazban. A nagyobb hiba azonban eredhet mérési problémából, hiányzó fontos jellemzőből, nem megfelelő modellből vagy a folyamat időközbeni változásából is. A hibát ezért először ki kell vizsgálnunk; a több adat önmagában nem minden problémát old meg. [20]

Ha az adatok elégtelen lefedettsége tűnik a fő oknak, célzott kísérlettel bővíthetjük a tanítóhalmazt. A szakértőkkel meghatározzuk, mely beállításokat és kombinációkat érdemes kipróbálni a technológiailag megengedett tartományon belül. A példánkban ilyen lehet néhány ritkábban előforduló átadási idő és hőkezelési beállítás vizsgálata. A kísérlet során a tényleges bemeneti jellemzőket és a végellenőrzési eredményeket is rögzítjük.

Nem minden bemeneti jellemző állítható közvetlenül. A gép beállításai befolyásolhatják például egy szenzorjel alakját, de a jelből számított szórást nem feltétlenül tudjuk előre meghatározott értékre állítani. A kísérleti tényezőket ezért különítsük el a később mért vagy számított jellemzőktől.

Az új adatok ellenőrzése után újrataníthatjuk a modellt, majd megvizsgáljuk, javult-e a teljesítménye a problémás tartományban és a szokásos gyártási körülmények között. A célzott kísérleti minta arányai eltérhetnek a napi gyártástól, ezért az azon mért összteljesítmény nem feltétlenül jellemzi a későbbi üzemi használatot. [20]

Ha a korábbi tesztteredmények már befolyásolták a fejlesztést és az adatgyűjtést, a végső értékeléshez új, független tesztadatokról is gondoskodnunk kell. Az új kísérleti adatokat tehát nem használhatjuk egyszerre teljes egészükben tanításra és a javulás független igazolására. [10]

Hasonlítsuk össze a csoportokat megtervezett körülmények között

Tervezett kísérletet a feltételezett csoportkülönbségek ellenőrzésére is végezhetünk. Tegyük fel, hogy két beszállító alapanyagából készülnek termékek, de az egyik beszállító sokkal több adattal szerepel a historikus állományban. Ráadásul előfordulhat, hogy a két alapanyagot eltérő gyártáson vagy más beállítások mellett használták.

Az eltérő mintaszám önmagában nem zárja ki a statisztikai összehasonlítást, és az azonos darabszám sem biztosítja automatikusan annak megalapozottságát. A kérdés az, hogy elegendő, megfelelően gyűjtött adatunk van-e a vizsgált különbség kimutatásához, és el tudjuk-e különíteni a beszállító szerepét más körülményektől. A szükséges mintaszámot a szakmailag fontos eltérés, a változékonyság és a kísérleti elrendezés alapján tervezzük meg. [12]

A példánkban mindkét alapanyaggal gyárthatunk összehasonlítható beállítások mellett. A gyártási sorrendet lehetőség szerint véletlenszerűsítjük, hogy az időbeli változások kevésbé kapcsolódjanak

egyetlen csoporthoz. Az ismert zavaró körülményeket, például a gyártási napot vagy gépet, blokkolással kezelhetjük: hasonló körülményeket képviselő blokkokon belül mindkét alapanyagot vizsgáljuk. [21]

A független ismétlések és a gyártási tételek szerepét is tisztáznunk kell. Egyetlen alapanyagból gyártott sok termék nem helyettesít több, egymástól független tételt, ha a beszállító általános hatására szeretnénk következtetni. Ha a gyártás miatt csak tételek vagy teljes gépbeállítások szintjén lehet változtatni, ehhez kell igazítanunk a kísérlet tervét és az elemzést. [21]

Vizsgáljuk az eltérő érzékenységet is

Előfordulhat, hogy az egyik alapanyagból készült termékek érzékenyebben reagálnak egy gyártási paraméter változására, mint a másik csoport. A példánkban a hosszabb átadási időhöz kapcsolódó minőségváltozás mértéke beszállítónként eltérhet. Ennek vizsgálatához mindkét alapanyagot több, összehasonlítható átadási idő mellett érdemes tesztelni.

Elegendő adat esetén külön modelleket is illeszthetünk a két csoportra, és összevethetjük azok becsléseit közös bemeneti tartományon. Másik lehetőség egy közös modell, amely a csoport és a gyártási paraméter közötti kölcsönhatást, más néven interakciót is leírja. Ilyenkor azt vizsgáljuk, hogy a paraméterhez kapcsolódó változás függ-e a csoporttól. [12]

A külön modellek eltérő jellemzőfontosságai önmagukban nem bizonyítanak eltérő fizikai érzékenységet. A becsült különbség bizonytalanságát, a modellek hibáját és az összehasonlított tartományt is figyelembe kell vennünk. Az ellenőrzött eredmények alapján a szakértő mérlegelheti, indokolt-e alapanyagoként eltérő gyártási beállítás. [19]

Építsünk az eredmények használatát támogató riportoló rendszert vagy alkalmazást

Egy adattudományi projekt eredménye többféle formában válhat használhatóvá. Készülhet belőle riport, prezentáció, oktatási anyag, munkautasítás, interaktív alkalmazás vagy más rendszerbe beépített prediktív modell. A megfelelő megoldást az határozza meg, kinek, milyen gyakran és milyen döntéshez van szüksége az eredményekre.

Amikor az elemzés és a megszerzett tudás a végtermék

Ha a cél egy meghatározott kérdés megválaszolása és az eredmény dokumentálása, nem feltétlenül szükséges folyamatos üzemi adatfeldolgozást kialakítanunk. A projekt végterméke lehet a modellel feltárt összefüggések értelmezése és az ezekből származó, ellenőrzött szakterületi tudás.

Az elemzésnek ilyenkor is ellenőrizhetőnek és lehetőség szerint megismételhetőnek kell lennie. Őrizzük meg az alkalmazott adatváltozatot vagy annak egyértelmű azonosítását, az adat-előkészítés és modellezés kódját, a beállításokat és a futtatási környezet leírását. A verziókövetés segít megállapítani, melyik eredmény pontosan milyen adatokból és eljárással készült. [22]

A riportban ismertessük a kérdést, az adatok eredetét, a kiértékelés módját, a metrikák jelentését és az eredmények korlátait. Legyen világos, mely megállapítás prediktív összefüggés, melyik ellenőrzendő hipotézis, és mit támasztanak alá további vizsgálatok. A gépi tanulás tudományos felhasználását, a reprodukálhatóságot és a jelentéskészítést részletesen tárgyalja Timo Freiesleben és Christoph Molnar *Supervised Machine Learning for Science* című könyve. [22]

Amikor a megoldást a mindennapi munkában használják

Ha a szakértő mérnökcsoporthoz rendszeresen használja a megoldást, az adatfeldolgozás és az alkalmazás működése is a fejlesztés része lesz. Az adatfeldolgozási folyamatnak, vagyis a pipeline-nak több

forrásból érkező adatokat kellhet összegyűjtenie és a kívánt gyakorisággal feldolgoznia. Ez lehet időzített frissítés vagy eseményhez kapcsolódó feldolgozás; nem minden feladathoz szükséges valós idejű működés.

Felhőszolgáltatásokkal helyi vállalati és felhőben tárolt adatok is összekapcsolhatók. Az Azure Data Factory például ütemezett adatmozgatási és átalakítási folyamatokat támogat, a helyi hálózat adatforrásainak eléréséhez pedig saját környezetben futó integrációs komponens használható. A kapcsolat megfelelő hálózati beállításokat és hozzáférési jogosultságokat igényel. [23]

A felhő használata önmagában nem teszi a rendszert robusztussá, és nem jelent minden felhasználónak fejlesztői hozzáférést. A működtetéshez, használathoz és módosításhoz szükséges jogosultságokat külön kell meghatározni. A megoldás helyi infrastruktúrán vagy a helyi és felhőszolgáltatások kombinációjával is megvalósítható.

A nyers adatok beérkezése után az automatizált feldolgozás frissíti a táblákat, ellenőrzi és tisztítja az adatokat, összekapcsolja a megfelelő termékhez tartozó méréseket, majd előállítja a modellhez szükséges jellemzőket. A bemenetek kialakításának összhangban kell maradnia a tanításkor alkalmazott eljárással. [11]

A robusztus működéshez azt is meg kell tervezni, mi történjen hiányzó, későn érkező vagy ismételt elküldött adatoknál. Egy sikertelen feldolgozás újraindítása például ne hozza létre ugyanazt a termékekordot még egyszer. A hibák legyenek észlelhetők és visszakereshetők; a hiányos adatból készült eredmény pedig ne jelenjen meg észrevétlenül teljes értékű becslésként.

Válasszunk a felhasználáshoz illő felületet

Ha a felhasználónak át kell tekintenie vagy vizsgálnia kell az eredményeket, grafikus felhasználói felületet (user interface, UI) alakíthatunk ki. Egy ilyen alkalmazásban statikus vagy interaktív ábrákat jeleníthetünk meg, szűrhetjük a termékeket és időszakokat, elemzéseket indíthatunk, illetve bemeneti értékeket adhatunk meg. A felületet a felhasználó tényleges döntéseihez és munkamenetéhez kell igazítanunk.

A modell üzemi alkalmazása azonban nem feltétlenül igényel önálló grafikus felületet. Az eredmény kerülhet adatbázisba, meglévő vállalati alkalmazásba vagy riasztási csatornára, illetve lekérhető programozási interfészen (API) keresztül. A mérnök vagy a kapcsolódó rendszer ezután az előre meghatározott szabályok szerint használhatja fel az előrejelzést.

A modell futtatása és újratanítása külön művelet. Ha ezek valamelyikét a felhasználó indítja, ehhez egyértelmű kezelési lehetőség szükséges, amely lehet a meglévő alkalmazás része is. Az újratanított modell nem válik automatikusan alkalmassá az üzemi használatra: teljesítményét ellenőrizni kell, és követhetővé kell tennünk, mikor melyik modellváltozat készítette az eredményeket.

A riport vagy alkalmazás elkészítésével tehát már konkrétan kipróbálható megoldásunk van. A következő lépésben a teljes működést teszteljük, az adat beérkezésétől a felhasználóhoz eljutó eredményig.

Teszteljük a teljes megoldást

A modell teljesítményének kiértékelése után a teljes megoldás működését is ellenőrizni kell. Az adat beérkezésétől a feldolgozáson és az előrejelzésen át a megjelenítésig vagy a kapcsolódó rendszernek küldött eredményig minden lépésnek megfelelően kell működnie. A tesztelés során hibákat tárhatunk fel, hiányosságokat fedezhetünk fel, és új funkciókra is javaslatot tehetünk.

Különítsük el a hibákat a hiányosságoktól és az új igényektől

Hibáról beszélünk, ha a megoldás eltér az elvárt működéstől. Előfordulhat például, hogy bizonyos adatoknál leáll az alkalmazás, nem jelenik meg a diagram, hibás a tengelyfelirat vagy a mértékegység, illetve elmarad az elemzés eredménye. Hiba lehet az indokolatlanul hosszú feldolgozási idő vagy az is, ha a rendszer rossz termékhez kapcsolja a helyesen kiszámított előrejelzést.

Egy téves előrejelzés önmagában még nem bizonyít szoftverhibát: a statisztikai modellnek akkor is van hibája, ha az alkalmazás megfelelően működik. A kivizsgálás során ezért különítsük el a modell pontatlanságát az adatok, a feldolgozás és a megjelenítés hibáitól.

Hiányosságot jelezhet a felhasználó, ha az alkalmazás működik, de egy valós munkafeladathoz nem ad elegendő információt vagy kezelési lehetőséget. Például látható a selejtarány növekedése, de nem kereshetők vissza az érintett termékazonosítók; megjelenik egy előrejelzés, de nem derül ki, mikor frissültek az alapjául szolgáló adatok; vagy hiányos mérések esetén nincs érthető tájékoztatás a felhasználó teendőiről.

Új igény lehet egy további szűrés beszállító vagy gyártási tétel szerint, egy új elemzés, magyarázó diagram vagy exportálási lehetőség. Hogy egy észrevétel hibajavításnak, hiánypótlásnak vagy új funkciónak számít-e, azt a korábban elfogadott követelmények alapján döntjük el. A szükséges módosításokat a felhasználókkal közösen rangsoroljuk.

Próbáljuk ki a valós munkameneteket

A fejlesztői ellenőrzéseket egészítsük ki felhasználói átvételi teszttel: a mérnökök a saját munkájukra jellemző feladatokat végezzenek el az alkalmazással. A példánkban válasszanak ki egy problémás gyártási időszakot, keressék vissza az érintett termékeket, és ellenőrizzék az átadási idő, valamint a végellenőrzési eredmények megjelenítését.

A szokásos adatok mellett próbáljuk ki az üres találati listát, a hiányzó és késve érkező méréseket, az ismételten feldolgozott rekordokat, valamint a nagyobb adatmennyiséget is. Vizsgáljuk meg, hogy a felhasználó a szerepkörének megfelelő adatokat és funkciókat éri-e el. Egy javítás után ellenőrizzük, hogy a korábban jól működő feladatok továbbra is helyes eredményt adnak-e.

Előre rögzítsük, mi alapján fogadjuk el a megoldást: például helyesen kapcsolja össze a termékadatokat, a frissítés a megállapodott időn belül befejeződik, és a mérnök önállóan végre tudja hajtani a kijelölt feladatot. A hibajegyzékben legyen reprodukálható a probléma: szerepeljen benne a bemenet vagy annak azonosítója, az alkalmazás verziója, az elvárt és a tapasztalt működés.

Ellenőrizzük az új modellváltozat bevezetését

Ha egy már üzemben működő modellt újratanítunk vagy másik modellre cserélünk, az offline kiértékelés után az üzemi környezetben is ellenőriznünk kell az új változatot. Ehhez több, egymással kombinálható bevezetési módszer áll rendelkezésre.

Árnyéküzemben (shadow testing vagy shadow deployment) az új modell az éles bemenetek másolatán fut, de az eredménye még nem befolyásolja a felhasználónak adott választ vagy a gyártási döntést. Így a régi és az új változat működése párhuzamosan megfigyelhető. [24]

Canary bevezetésnél az új változat kezdetben az éles kérések kis részét szolgálja ki. Megfelelő eredmények esetén ez az arány fokozatosan növelhető. Blue-green bevezetésnél két elkülönített környezetet vagy telepítést tartunk fenn: az egyik szolgálja ki az éles működést, miközben a másikban előkészítjük és ellenőrizzük az új változatot. Ezután átirányíthatjuk rá a forgalmat. A két környezet közötti átállás fokozatos forgalommegosztással is megvalósítható. [24]

A Databricks Model Serving például támogatja több modell vagy modellverzió kiszolgálását egy végponton, és a kérések százalékos megosztását közöttük. Ez lehetőséget ad a fokozatos átállás kialakítására. A forgalom megosztása azonban nem azonos ugyanannak a kérésnek a mindkét modellhez történő elküldésével. [25]

A gyártási példánkban az új modell előrejelzéseit először csak naplózhatjuk, majd a termékazonosító alapján összekapcsolhatjuk őket a később elkészülő végellenőrzési eredményekkel. A gyors és hibamentes futás még nem igazolja a prediktív teljesítményt: ehhez a tényleges célértékek is szükségesek. Ha a végellenőrzés eredménye késik, a teljesítményértékelés is csak később végezhető el.

Az átállás előtt határozzuk meg, milyen eredmények mellett léphet tovább a bevezetés, és mikor kell visszatérni a korábbi változathoz. A visszaállítás lehetőségét a modellhez tartozó adat-előkészítéssel együtt biztosítsuk. A már végrehajtott gyártási beavatkozásokat egy szoftveres visszaállítás nem vonja vissza, ezért a próbaüzem döntési jogosultságait is egyértelműen rögzíteni kell.

Illesszük be a megoldást a mérnöki munkafolyamatokba

Egy ipari adattudományi megoldás akkor teremt tartós értéket, ha a mérnökök a munkájuk során fel tudják használni az eredményeit. A technikailag elkészült és tesztelt alkalmazás önmagában még nem biztosítja, hogy megváltoznak a döntések vagy javul a folyamat.

A szakértői csapatnak értenie kell, milyen adatokból dolgozik a megoldás, mire használható, milyen körülmények között bizonytalan, és milyen teendő következik az eredményeiből. Ehhez nem feltétlenül kell a modell minden matematikai részletét ismernie. A modell magyarázata ugyanakkor nem helyettesíti a fizikai mechanizmus igazolását: egy jól előrejelző összefüggés továbbra is lehet ellenőrzendő hipotézis. [19]

A fejlesztés közbeni rendszeres egyeztetés segít abban, hogy a megoldás illeszkedjen a tényleges igényekhez. Az elfogadást azonban a követelmények teljesülésével és a felhasználói tesztek eredményeivel is ellenőriznünk kell. A jó együttműködés önmagában nem garantálja, hogy minden szükséges funkció elkészült vagy minden munkakörülményre gondoltunk.

A bevezetés során rögzítsük, ki, mikor és milyen döntéshez használja az eredményt. Legyen felelőse a jelzések értékelésének, a szükséges beavatkozásnak és a működési hibák kezelésének. Az új feladatokat építsük be a meglévő munkamenetbe, szükség szerint munkautasítással, rövid oktatással és közösen végigvitt példákkal.

A gyártási példánkban megállapodhatunk abban, hogy a műszakért felelős mérnök áttekinti a hosszú átadási idővel érintett termékeket és a hozzájuk tartozó előrejelzéseket. Egy jelzéshez meghatározott ellenőrzés vagy beavatkozási eljárás tartozik. Így a diagram vagy predikció egy konkrét mérnöki feladat részeként jelenik meg.

Ha a felhasználók mégsem alkalmazzák a megoldást, vagy figyelmen kívül hagyják az eredményeket, kezdeményezzünk párbeszédet. Lehet, hogy túl későn érkezik a jelzés, sok a téves riasztás, nehézkes a felület, nem világos a teendő, vagy a felhasználónak nincs lehetősége beavatkozni. Az is kiderülhet, hogy a megoldás haszna elmarad a várakozásoktól. A visszajelzések alapján az alkalmazást, a munkafolyamatot vagy az eredeti célkitűzést is módosíthatjuk.

Az üzemi használat megkezdésével új szakasz indul: követnünk kell a megoldás működését, a felhasználói tapasztalatokat és a ténylegesen elért eredményeket. Ezek adják a későbbi fejlesztés alapját.

Kövessük nyomon a működését és a hatását

A fejlesztési folyamat az átadással nem ér véget. Az alkalmazás vagy modell üzemi működését és gyakorlati hatását is követnünk kell. A nyomon követés célja, hogy időben észleljük a problémákat, megértsük a használat tapasztalatait, és ellenőrizzük, megvalósul-e a várt értékteremtés.

Figyeljük a működést és a felhasználást

Egy diagramokat, statisztikai elemzéseket és előrejelzéseket tartalmazó alkalmazásnál naplózhatjuk a futtatások számát, gyakoriságát és időbeli eloszlását, a feldolgozási időt, valamint a hibákat. Az adatok frissességét is érdemes követni: az alkalmazás látszólag működhet, miközben egy elakadt adatbetöltés miatt elavult eredményeket mutat.

A naplókából rendszeres összesítéseket és időszakos jelentéseket készíthetünk. A működést veszélyeztető eltérésekhez gyorsabb jelzést és kijelölt felelőst rendelhetünk. A futtatásszám értelmezésénél különítsük el az automatikus háttérfolyamatokat a tényleges felhasználói műveletektől: a sok futás önmagában nem bizonyítja, hogy a megoldást döntésekhez is használják.

Mérjük a modell teljesítményét és a gyakorlati eredményt

A prediktív teljesítményt az előrejelzések és a később megismert tényleges eredmények összevetésével követhetjük. A modellváltozatot és a predikció időpontját is rögzítsük. Osztályozásnál például a NOK-osztály recall és precision értékét, regressziónál az abszolút hibát figyelhetjük. A teljes állomány mellett vizsgáljuk a fontos termékváltozatokat, gyártósortokat és időszakokat is. A metrikák mellett tüntessük fel az értékelt esetek számát: néhány NOK-termékből számított recall erősen ingadozhat. [26]

Kérjünk visszajelzést a felhasználóktól arról, hogy csökkent-e a kézi adatelőkészítés és riportkészítés ideje, gyorsabbá vált-e a problémák felismerése, illetve javult-e a gyártási folyamat. A visszajelzéseket lehetőség szerint egészítsük ki mérhető mutatókkal: selejtarány, utómunka, energiafelhasználás vagy az elemzéshez szükséges munkaidő.

A változást a korábban rögzített kiindulási állapothoz és összehasonlítható gyártási körülményekhez viszonyítsuk. A kisebb selejtarányt más változás, például új alapanyag vagy karbantartás is okozhatja. A bevezetés előtti és utáni különbséget ezért nem tulajdoníthatjuk automatikusan a modellnek. A gazdasági eredménynél a megtakarítás mellett a működtetés és a jelzések nyomán végzett beavatkozások költségét is vegyük figyelembe.

A modell használata magát a megfigyelhető eredményt is megváltoztathatja. Ha egy NOK-előrejelzés miatt a mérnök időben korrigálja a folyamatot, a termék végül megfelelő lehet. Ez nem feltétlenül téves riasztás. Ha pedig a terméket korábban kivonják a folyamatból, a szokásos végteszteredmény hiányozhat. A beavatkozásokat és a hiányzó végellenőrzések okát ezért szintén rögzítsük, és az értékelési tervben kezeljük ezeket az eseteket.

Vizsgáljuk ki a teljesítményromlás okait

Az üzemi teljesítmény idővel romolhat, vagy már az első alkalmazáskor elmaradhat a fejlesztés során mért értéktől. Többféle okot érdemes elkülönítenünk:

Bemeneti adateloszlás változása (data drift): a modellhez érkező jellemzők eloszlása eltér a referencia-időszakétól. Például más arányban gyártunk termékváltozatokat, vagy gyakoribbá válik a hosszú átadási idő. Az eltolódás figyelmeztető jel, de önmagában nem bizonyít teljesítményromlást. [26, 27]

A bemenet és a célérték kapcsolatának változása (concept drift): azonos megfigyelt bemenetek mellett megváltozik a célérték feltételes eloszlása. A példánkban egy technológiai módosítás után ugyanakkora

átadási idő és azonos rögzített beállítások mellett más végteszteredmények jellemzők. A bemeneti eloszlás pusztá figyelése nem elegendő a változás igazolására. [27]

Adatminőségi vagy feldolgozási hiba: módosul az adatséma, kimarad egy mérés, megváltozik a mértékegység, vagy másképp számítjuk a jellemzőt, mint tanításkor. Ilyenkor az adat útját és az előfeldolgozást kell kivizsgálnunk. [11, 26]

Elégtelen általánosítás: a tanítóadatok nem fedték le megfelelően az üzemi eseteket, vagy a korábbi kiértékelés túl kedvező képet adott. Ez nem feltétlenül időbeli változás: egy már meglévő modellkorlát is most válhat láthatóvá. [10, 20]

A vizsgálat eredménye alapján javíthatjuk az adatfeldolgozást, bővíthetjük a méréseket, módosíthatjuk a modellt vagy újrataníthatjuk friss adatokon. Az újratanítás nem minden hibára megfelelő válasz: egy hibás mértékegységet vagy rossz termék-hozzárendelést először az adatoldalon kell helyrehozni. Az új változatot a korábban bemutatott kiértékelési és bevezetési lépéseken keresztül alkalmazzuk.

A tapasztalatok alapján fejlesszük tovább

Az adattudományi megoldások fejlesztése a működés közben szerzett tapasztalatokra épül. A következő módosítást indokolhatja felhasználói igény, feltárt hiba, új mérési lehetőség vagy a technológia változása. A fejlesztés célját minden alkalommal konkrétan fogalmazzuk meg: milyen problémát oldunk meg, és miből fogjuk látni a javulást?

A változtatások ellenőrzött bevezetését a szoftverfejlesztésből ismert CI/CD gyakorlatok támogatják. A folyamatos integráció (continuous integration, CI) során a kódmódosításokat rendszeresen egyesítjük, és automatizált ellenőrzéseket futtatunk. A folyamatos szállítás (continuous delivery, CD) a tesztelt változatot kiadásra kész állapotba hozza; az élesítés külön döntés maradhat. A continuous deployment szintén CD rövidítéssel szerepelhet: ennél a sikeres ellenőrzések után az élesítés is automatikus.

Gépi tanulásnál a kód mellett az adatfeldolgozás és a modell ellenőrzése is fontos. A folyamatos tanítás (continuous training, CT) külön fogalom: az újratanítási folyamat időzítve vagy meghatározott esemény hatására indulhat. Ez nem jelent minden új adatpont után történő tanítást, és nem helyettesíti az új modell értékelését. A CI/CD és CT szerepét a Google Cloud MLOps-útmutatója részletesen tárgyalja. [28]

Gyártóüzemben, földtudományi vizsgálatban vagy vegyi laboratóriumban egy új szenzor korábban nem megfigyelt folyamatjellemzőt tehet mérhetővé. A példánkban a hőkezelő állomás előtti közvetlen hőmérsékletmérés segíthet megvizsgálni, mennyire jó helyettesítő jellemző volt az átadási idő a termék hőállapotának leírására.

Az új mérés hasznát azonban ellenőriznünk kell. Tisztázzuk a pontosságát, a termékhez rendelését és azt, hogy az előrejelzés időpontjában már rendelkezésre áll-e. Be kell építenünk az adatgyűjtésbe és az előfeldolgozásba; szükség lehet új jellemzőkre, statisztikai elemzésekre vagy új modell tanítására. Ha a korábbi termékeknél még nem volt ilyen mérés, előbb elegendő új adatot kell gyűjtenünk az összehasonlításhoz.

A módosítás után ismét kiértékeljük a modellt, teszteljük a teljes megoldást, beillesztjük a munkafolyamatokba, majd követjük a működését és a hatását. A tapasztalatok új kérdéseket vethetnek fel, amelyek visszavezetnek a folyamat megértéséhez vagy a mérések megtervezéséhez. Így a tizennégy lépés egymásra épülő, visszacsatolásokkal működő fejlesztési folyamattá válik.

Az adattudomány és a tudomány jövője az AI korában

A generatív mesterséges intelligencia (generative AI, GenAI) eszközei a mérnöki és kutatói munka számos részében használhatók: segíthetnek a programozásban, az információkeresésben és -kivonatolásban, a szövegalkotásban, valamint matematikai és fizikai modellek megértésében és fejlesztésében. Kódvázlatot, magyarázatot vagy megoldási javaslatot készíthetnek, amelyből kiindulva tovább dolgozhatunk.

Egyes részfeladatok elvégzését jelentősen felgyorsíthatják. A tényleges időnyereség azonban függ a feladattól, az eszköztől és a felhasználó tudásától, valamint attól, mennyi ellenőrzést és javítást igényel az eredmény. Egy gyorsan elkészített program akkor jelent előrelépést, ha a megfelelő problémát oldja meg, és megbízhatóan működik.

A gyors eredmény mellett érdemes figyelni arra is, milyen tudás keletkezik a munka során. Terence Tao matematikus egy 2026-os előadásában hangsúlyozta, hogy egy matematikai eredmény előállítása és ellenőrzése mellett annak magyarázata, megértése és közösségi felhasználhatósága is lényeges. A Simons Foundation beszámolója szerint Tao a megoldáshoz vezető út bemutatását és a tudás átadhatóságát is kiemelte. [29]

A Big Think videójában Tao a lassabb problémamegoldás értékére is felhívja a figyelmet:

“There is a value doing things a slow way” ... “Learn a lot of extra insight beyond just getting the answer they are trying to seek. And they can find they can discover new phenomena”.[30]

Ez a gondolat a mérnöki adattudományra is alkalmazható. A nehéz problémán való gondolkodás, a hipotézisek vizsgálata és az elsőre sikertelen próbálkozások során olyan kapcsolatokat vagy jelenségeket fedezhetünk fel, amelyekre a munka kezdetén még nem számítottunk. Egy kész válasz önmagában nem feltétlenül adja át ezeket a felismeréseket.

A gyártási példánkban egy AI-eszköz gyorsan elkészítheti az átadási idő és a végteszteredmény kapcsolatát bemutató ábrát. A mérnöki vizsgálat során azonban felmerülhet, hogy a kapcsolat csak bizonyos alapanyagoknál jelenik meg, egy gépállapothoz kötődik, vagy a két állomás időbélyegeinek eltérése torzítja. Ezek tisztázása közben az eredeti kérdésünk is változhat: lehet, hogy új mérésre vagy másképpen megtervezett kísérletre van szükségünk.

Az információ értelmezéséhez és gyakorlati felhasználásához tudnunk kell, miért tettük fel a kérdést, milyen feltételek mellett érvényes a válasz, és milyen döntéshez akarjuk felhasználni. A lassabb, elmélyült vizsgálódás értéke ebben a megértésben rejlik. Nem minden részfeladaton szükséges hosszasan dolgoznunk, de érdemes időt hagynunk a feltevések ellenőrzésére és a váratlan eredmények értelmezésére.

Az AI ezt az elmélyülést is támogathatja. Kérhetünk tőle alternatív magyarázatokat, ellenpéldákat, összehasonlítható modelleket vagy a feltevéseinket próbára tevő kérdéseket. A javaslatokat ezután adatokkal, számításokkal, szakirodalommal és szükség esetén kísérletekkel ellenőrizzük. A rutinmunka automatizálásával felszabaduló időt így a probléma jobb megértésére fordíthatjuk.

A kísérletezés és a megoldáshoz vezető út megértése új kutatási irányokat is kijelölhet. Ha azt is el tudjuk mondani, miért működik egy megoldás, hol vannak a korlátai, és mely próbálkozások vezettek hozzá, mások könnyebben építhetnek az eredményeinkre. Ez a tudásmegosztás a mérnökök, kutatók és adattudósok együttműködését is segíti.

A jegyzet célja ezért az is, hogy a segítsen hallgatónak kérdéseket megfogalmazni, a kapott válaszokat megvizsgálni és mérnöki jelentésüket megérteni. Az AI-eszközök fejlődésével ezek a képességek

segíthetnek abban, hogy a gyorsabban elérhető eredményekből ellenőrzött tudás és használható megoldás szülessen.

Hivatkozások

A számozott hivatkozások a módszertani részek forrásait jelölik. A gyártási példa és a projektmenet a szerző saját szemléltetése. Az [1–12] online források ellenőrzésének dátuma: 2026. szeptember 13.; a [13–19] forrásoké: 2026. szeptember 14. A [20–23] források ellenőrzésének dátuma: 2026. szeptember 15. A [24–25] források ellenőrzésének dátuma: 2026. szeptember 15. A [26–28] források ellenőrzésének dátuma: 2026. szeptember 16. A [29] forrás ellenőrzésének dátuma: 2026. szeptember 17.

[1] Christ, M., Braun, N., Neuffer, J. és Kempa-Liehr, A. W. (2018). Time Series Feature Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package). Neurocomputing, 307, 72–77. DOI: 10.1016/j.neucom.2018.03.067. A módszer használatához ellenőrzött forrás: tsfresh dokumentáció, Introduction. [Dokumentáció](#) · [DOI](#)

[2] Kuhn, M. és Johnson, K. (2019). Feature Engineering and Selection: A Practical Approach for Predictive Models. Chapman & Hall/CRC. Online szerzői változat, Preface. [Online könyv](#)

[3] Molnar, C. (2022). Modeling Mindsets: The Many Cultures of Learning From Data. A különböző modellezési szemléletek áttekintéséhez; ellenőrzött szerzői ismertető és tartalomjegyzék. [Szerzői könyvoldal](#)

[4] NIST/SEMATECH. e-Handbook of Statistical Methods. 1.3.5 Quantitative Techniques; 1.3.5.3 Two-Sample t-Test for Equal Means. [Hipotézisvizsgálat](#) · [Kétmintás t-próba](#)

[5] Scikit-learn fejlesztői. Getting Started. A felügyelt tanulás és a becslési feladatok bevezetése. [Dokumentáció](#)

[6] Scikit-learn fejlesztői. Clustering. Különösen a klaszterezési módszerek áttekintése és az új megfigyelések hozzárendelésének lehetőségei. [Dokumentáció](#)

[7] Shmueli, G. (2010). To Explain or to Predict? Statistical Science, 25(3), 289–310. DOI: 10.1214/10-STS330. [Szerzői tanulmány](#)

[8] PyMC fejlesztői. Introductory Overview of PyMC. Különösen A Motivating Example: Linear Regression és Model Specification. [Bayes-i regressziós példa](#)

[9] Box, G. E. P. (1979). Robustness in the Strategy of Scientific Model Building. In R. L. Launer és G. N. Wilkinson (szerk.), Robustness in Statistics, 201–236. Academic Press. Az idézett mondat a 202. oldalon szerepel. [Eredeti tanulmány digitalizált másolata](#)

[10] Scikit-learn fejlesztői. Cross-validation: evaluating estimator performance. Tanító-, validációs és tesztadatok; időbeli és csoportos felosztások. [Dokumentáció](#)

[11] Scikit-learn fejlesztői. Common pitfalls and recommended practices. Data leakage. [Információszivárgás](#)

[12] NIST/SEMATECH. e-Handbook of Statistical Methods. 5.3 Choosing an experimental design. [Kísérlettervezés](#)

[13] Scikit-learn fejlesztői. Metrics and scoring: quantifying the quality of predictions. Klasszifikációs metrikák. Online forrás

[14] Scikit-learn fejlesztői. Post-tuning the decision threshold for cost-sensitive learning. Online forrás

[15] Scikit-learn fejlesztői. Tuning the decision threshold for class prediction. Online forrás

[16] Molnar, C. Interpretable Machine Learning. Online könyv. Online forrás

- [17] Scikit-learn fejlesztői. `r2_score`. Online forrás
- [18] Molnar, C. Interpretable Machine Learning. SHAP című fejezet. Online forrás
- [19] SHAP fejlesztői. Be careful when interpreting predictive models in search of causal insights. Online forrás
- [20] Freiesleben, T. és Molnar, C. (2024). Supervised Machine Learning for Science: How to stop worrying and love your black box. 7. fejezet, Generalization. [Online forrás](#)
- [21] NIST/SEMATECH. e-Handbook of Statistical Methods. Completely randomized designs; Randomized block designs; How can I account for nested variation. [Online forrás](#) · [Randomizálás](#) · [Korlátozott randomizálás](#)
- [22] Freiesleben, T. és Molnar, C. (2024). Supervised Machine Learning for Science. 13. fejezet, Reproducibility; 14. fejezet, Reporting. [Online forrás](#) · [Reporting](#)
- [23] Microsoft. Introduction to Azure Data Factory; Create a self-hosted integration runtime. [Online forrás](#) · [Áttekintés](#)
- [24] Microsoft. Perform safe rollout of new deployments for real-time inference. Azure Machine Learning. [Online forrás](#)
- [25] Databricks. Create custom model serving endpoints. Több modellváltozat kiszolgálása és a forgalom megosztása. [Online forrás](#)
- [26] Microsoft. Azure Machine Learning model monitoring. [Online forrás](#)
- [27] Evidently AI. What is concept drift in ML, and how to detect and address it. [Online forrás](#)
- [28] Google Cloud. MLOps: Continuous delivery and automation pipelines in machine learning. [Online forrás](#)
- [29] Beaufore, J. (2026. augusztus 13.). Watch: Fields Medalist Terence Tao on Artificial Intelligence and Why We Do Math. Simons Foundation. Beszámoló Tao Mathematics in the Age of AI című előadásáról. [Online forrás](#)
- [30] The paradox at the heart of AI and science | Terence Tao. [Online forrás](#)

További olvasmányok

- James, G., Witten, D., Hastie, T., Tibshirani, R. és Taylor, J. (2023). An Introduction to Statistical Learning with Applications in Python. Springer. A regresszió, osztályozás, újramintavételezés és felügyelet nélküli tanulás részletes tanulásához. [A könyv szerzői oldala](#)
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. és Rubin, D. B. (2013). Bayesian Data Analysis. 3. kiadás. Chapman & Hall/CRC. A Bayes-i módszerek elmélyítéséhez. [A könyv szerzői oldala](#)