

SZAKDOLGOZAT

Horváth Tomáš

2024



Magyar Agrár- és Élettudományi Egyetem

Budai Campus

Élelmiszertudományi és Technológiai Intézet

Pálinkamester szaktanácsadó szakirányú továbbképzési szak

SZENZOROK ALKALMAZÁSI LEHETŐSÉGEI AZ ERJEDÉS MONITOROZÁSÁHOZ

SZAKDOLGOZAT

Konzulensek: Dr. Kun Szilárd
Egyetemi docens
Biomérnök és Erjedésipari
Technológia Tanszék

Dr. Sipos László
Egyetemi docens
Árukezelés, Kereskedelem,
Ellátási Lánc és Érzékszervi
Minősítési Tanszék

Készítette: Horváth Tomáš

2024

Tartalomjegyzék

1. Bevezetés és célkitűzések	1
2. Szakirodalmi áttekintés	4
2.1. Paraméterek, szenzorok és adattípusok	4
2.2. Adatelemzési módszerek	7
2.2.1. Adatbányászat	8
2.2.2. Gépi tanulás	11
2.3. „Survey” cikkek	16
2.3.1. Erjesztés monitorozása a bio- és gyógyszeriparban	16
2.3.2. e-Orr és e-Nyelv alkalmazása az élelmiszeriparban	17
2.3.3. e-Orr és e-Nyelv alkalmazása a szesziparban	20
2.4. Alkoholos erjesztés in-line és on-line monitorozása	23
2.4.1. Szín és fenoltartalom	23
2.4.2. Szárazanyagtartalom	24
2.4.3. Szén-dioxid	25
2.4.4. Hőmérséklet	25
2.4.5. Savtartalom és pH	26
2.4.6. Alkoholtartalom	26
2.4.7. Aromaanyagok	27
3. Alkalmazott módszerek	29
4. Eredmények és értékelésük	31
4.1. Értékelés szakmai szempontok alapján	31
4.2. Értékelés kutatói szemszögből	33
4.2.1. Adatok elemzése	34
5. Következtetések és javaslatok	37
6. Összefoglalás	39
Irodalomjegyzék	41
Ábrák jegyzéke	52
Táblázatok jegyzéke	53
Hallgatói nyilatkozat	54
Konzulensek nyilatkozata	55

1. fejezet

Bevezetés és célkitűzések

Irányított erjesztés (IE) nélkül ma már nemigen beszélhetünk modern pálinka- és párlatkészítésről (Panyik, 2013), illetve szeszgyártásról¹ (Jacques és tsai., 2003). IE alatt arra törekszünk, hogy a felhasznált fajlesztő törzs számára optimális körülményeket biztosítsuk a cefrében (úm. hőmérséklet; pH; szárazanyagtartalom; fizikai-, vegyi- és mikrobiológiai „tisztaság”) a teljes erjedési folyamat, valamint, ha szükséges, a cefre tárolásának ideje alatt, a minél jobb eredmény (pl. jó aroma- vagy alkoholkihozatal) elérésének érdekében.

Az IE főbb elemeinek fontosságát, mint pl. a fajlesztő használatát, a cefre „savasságának” beállítását, vagy a cefre megfelelő hőmérsékleten tartását, egyébként már a múlt század elején is felismerték. Hérics-Tóth és Osztróvszky (1918) írják, az „Az erjedés lefolyása és ellenőrzése” c. fejezetben, hogy *„A tudatosan dolgozó nemcsak az erjesztő helyiségnek, de a czefréknek hőmérsékletét is figyelemmel kíséri. E célból a hordókba, erjesztőedényekbe hőmérőt illeszt. (Nagyobb üzemeknél az erjesztőedényekbe oly csővezetékeket szerelnek fel, amelyekben keresztül, hideg és ellenkezőleg meleg vizet lehet keresztül áramoltatni. Az erjedés ily módon tetszés szerinti hőfoknál fog végbemenni”*.

Az IE paraméterei közül a hőmérséklet talán az egyetlen, amit folyamatosan, beavatkozás nélkül, egyszerűen és olcsón tudunk monitorozni, elég egy hőmérővel ellátni az erjesztőtartályt. Az erjedés további fontos paraméterei (mint pl. a pH, szárazanyagtartalom, alkoholtartalom, illósav tartalom) laboratóriumi műszeres vizsgálatokkal mérhetők (Ujszászi és Sólyom, 1980). Ezen méréseknek egy része már viszonylag egyszerűen kezelhető, könnyen és olcsón elérhető műszerekkel elvégezhető, amelyek kezeléséhez azonban bizonyos szintű ismeretek szükségesek. A kereskedelmi szeszfőzdék többsége rendelkezik is kisebb-nagyobb felszereltségű saját laboratóriummal, de a minőségi pálinkát előállító magánfőzőknél is már elengedhetetlen kellékként szerepelnek ezek a műszerek. Az emberi munkaerő, illetve a mérések elvégzéséhez szükséges szaktudás viszont még mindig nem megkerülhető, és – bármennyire jól felszerelt is egy ilyen laboratórium – a mérések nem vagy csak nehezen automatizálhatók.

¹Mivel a dolgozat témájának szempontjából mindegy, hogy pálinka-, párlat- vagy mezőgazdasági szeszgyártáshoz erjesztünk, a továbbiakban ezeket az általános „szesz(főzde)” gyűjtőfogalom helyettesíti.

Az informatika és az informatikai eszközök fejlődése lehetővé tette a szeszgyártási folyamatok, főként a lepárlás monitorozását és automatizálását (Offner, 2018), laboratóriumi (Enyedi és tsai., 2022) és ipari méretekben (Srinivasan és tsai., 2013) egyaránt. Itt a lepárló berendezés különböző pontjain mért hőmérséklet a legfontosabb paraméter, e mellett még a deflegmátorba, esetleg a Pistorius tányérokra engedett hűtővíz átfolyási sebessége, a fűtési teljesítmény és (ahol a technológiai megoldás engedi) a tányérokra lévő reflux magassága kerül monitorozásra. Ezen paraméterek monitorozása alapján a lepárlást vezérelni lehet, ami a professzionális lepárlóknál már természetesnek számít.

Az IE automatizáltságát illetően viszont még mindig azt látjuk, hogy a legtöbb piacon elérhető, széles körben használt erjesztő tartály csak a cefre hőmérsékletét monitorozza folyamatosan, ami alapján vezérli a tartály hűtését/fűtését. A cefre többi paramétereit, ahogyan fent említettem, pedig már manuálisan kell ellenőriznie a főzőmesternek. A könnyen és olcsón elérhető szenzorok sokszínűsége alapján természetesnek tűnik, hogy megvizsgáljuk alkalmazhatóságukat az erjesztés monitorozásánál is.

1. Kutatási kérdés: Hogyan monitorozzák az erjesztést más területeken (pl. a bio- és az élelmiszeriparban), annak milyen paramétereit és milyen típusú szenzorokkal mérik?

Különös figyelmet érdemel az élelmiszeriparban egyre gyakrabban használt elektronikus orr (e-Orr) és elektronikus nyelv (e-Nyelv) alkalmazhatóságának vizsgálata az erjesztés monitorozásához a szeszgyártásnál is, amiket a borászatban (Lozano és tsai., 2016) és a sörgyártásban (Toko és Tahara, 2016) már alkalmaznak. Az e-Orr és e-Nyelv segítségével – elvileg – lehetséges néhány aroma- és illatkomponens meglétét mérni a cefrében. A cefre azon komponenseinek korai felismerése lehet különösen értékes a főzőmester számára, amelyek a hibás erjedés indikátorai. Ilyenek lehetnek a nem kívánatos mikroorganizmusok metabolikus melléktermékei, mint pl. az ecetsav, vajsav, hidrogén-cianid, akrolein, stb.

2. Kutatási kérdés: Milyen feladatoknál és milyen anyagok felismerésére/mérésére használják az elektronikus orrot és az elektronikus nyelvet a szesziparban?

Az erjesztés a szeszgyártási folyamatoknak csak az egyik lépése, így feltételezem, hogy a dolgozatom témájával, az erjesztés monitorozásával is kevesebben foglalkoznak.

3. Kutatási kérdés: Milyen szenzorokkal, milyen módszerekkel és a cefrének milyen paramétereit monitorozzák az alkoholos erjesztésnél?

A szenzorok alkalmazhatóságának vizsgálatánál még azt is szem előtt kell tartani, hogy az általuk gyűjtött adatok milyen nagyságrendűek, hol és milyen gyakorisággal mérünk illetve, hogy ezeket az adathalmazokat milyen módszerekkel lehet elemezni, az általános statisztikai eljárásokon kívül.

4. Kutatási kérdés: Milyen adatbányászati és gépi tanulási módszerek alkalmazhatóak a szenzoradatokra, az alkalmazásuknak milyen előnyök és hátrányai vannak?

Az adatok elemzésének lehetőségeit azért fontos elemezni, hogy azokne csak az adott erjesztési folyamat(ok)-ról adjanak információt a főzőmesternek, hanem az adatokban rejlő összefüggések és mintázatok az IE témakörében végzett kutatásokat is támogassák.

A dolgozat célja: Dolgozatom áttekintő jellegű. Célja egy olyan, koherens információ- és tudásbázis létrehozása, amely jó kiindulási támpontot ad egy komplex, moduláris alkoholos erjesztést monitorozó rendszer fejlesztéséhez, beleértve a különböző szenzorok és adatelemzési módszerek előnyeinek és hátrányainak vizsgálatát és elemzését.

Mint általában, az adott témában is rengeteg információ található az interneten különböző, szeszgyártáshoz kapcsolódó fórumokon és blogokon. Ezek többnyire ún. „garázs” projektek vagy pedig kommersz termékek leírásai. Az előbbi esetben többnyire pontatlan dokumentációkkal és/vagy szubjektív véleményeken alapuló elemzésekkel találkozunk. Az utóbbi esetben pedig már készen kapható termékek, technikai részleteket többnyire felületesen bemutató marketing leírásai tölthetők le. Az ilyen típusú információk esetében azzal is számolni kell, hogy ezek nem biztos, hogy idővel elérhetők lesznek (a felhasználó törli a bejegyzéseit a blogról, az adott terméket kivonja a cég a forgalomból, stb.). Ezért, vannak a dolgozat fókuszában a *tudományos publikációk*. Itt érdemes megjegyezni, hogy a szakirodalomban sok olyan cikk jelenik meg, amelyek alapján nehéz az adott kutatást reprodukálni, illetve a bemutatott kísérletek és következtetések sem hibáktól mentesek vagy objektívek. Kutatásunk kitér ezen faktorok elemzésére is. Mivel a tudományos publikációk általában átesnek valamilyen szintű bírálati folyamaton, így feltételezzük, vagy inkább megelőlegezzük az ezekben bemutatott kutatások hitelességét.

A következő, 2. fejezetben az első, a második és a harmadik kutatási kérdésre összontosítok, a szakirodalom feltérképezésével. A fejezet elején bemutatom az erjesztés illetve a cefre azon a paramétereit is amiket érdemes (lenne) monitorozni és a különböző szenzor típusokat. Továbbá, röviden áttekintem a lehetséges adattudományi és gépi tanulási módszereket. A 3. fejezetben röviden leírom az irodalom feldolgozásának, általam választott menetét. A 4. fejezetben összefoglalom az irodalmi áttekintésből szerzett tudnivalókat és elemzem a különböző fejlesztési irányokat, különös tekintettel a pálinkafőzés sajátosságaira. A 5. fejezetben felvázolom egy általam javasolt, moduláris monitorozó rendszer felépítését.

A pálinkakészítés legtöbb lépése mára már jól automatizált, gondoljunk csak a modern gyümölcsfeldolgozó gépekre vagy a digitálisan vezérelhető lepárló berendezésekre. A folyamat egyik legkritikusabb lépésének, az erjesztésnek az egyik fontos eleme a monitorozás, ami viszont még mindig kézi erőt igényel. Ennek automatizálásához, még sok kutatásra és fejlesztésre van szükség, ahhoz, hogy nem csak a laboratóriumban, hanem a gyakorlatban is használható termék, eszköztár jöjjön létre. Dolgozatom ezt a kutatás-fejlesztési munkát hivatott támogatni, amin magam is dolgozni szeretnék a jövőben.

2. fejezet

Szakirodalmi áttekintés

2.1. Paraméterek, szenzorok és adattípusok

Ebben a fejezetben az erjesztés és a cefre főbb paramétereit, és az azokat monitorozni képes szenzortípusokat vázoljuk, röviden ismertetve azok alapelveit, előadás anyagok (Gillay, 2023; Abrankó, 2023) és internetes források (termékleírások, vitaforumok és blogok) alapján. Konkrét szenzor típusokra, termékekre ez a fejezet nem tér ki, ezek – amennyiben az elemzett cikkekben le vannak írva – a 2.4. fejezetben említésre kerülnek.

Hőmérséklet A cefre hőmérsékletét legjobban közvetlenül, a hőmérőt a cefrébe belemerítve mérhetjük. A különböző működési elvű hőmérők közül az *ellenállásos hőmérő* a megfelelő erre a célra, ami különböző fémeknek a hőmérséklet függvényében változó villamos ellenállásának mérésén alapszik, és három típusát ismerjük: i) a fém ellenállásos hőmérő a legpontosabb, de viszonylag lassan érzékeli a változásokat és drága; ii) a termoelemes hőmérő olcsó, de a hosszabb csatlakozó vezetékben fellépő ellenállás zajossá teheti a mérést; iii) a félvezetős hőmérő széles körben használt, olcsó, kevésbé pontos mint a fém ellenállásos de pontosabb mint a termoelemes hőmérő¹.

Szén-dioxid (CO₂) Az erjedési folyamat előrehaladásának és állapotának egyik fontos indikátora az élesztők által termelt CO₂ mennyisége, amit gázszenzorok segítségével tudunk mérni. Ezek legnépszerűbb típusa az *elektrokémiai* alapon működő gázszenzor. Ebben egy olyan érzékelő anyag van (pl. ón-dioxid), aminek a vezetőképessége a levegő szennyezettségével arányosan nő, amit egyébként a hőmérséklet is befolyásol. Ezeket a szenzorokat *MQ* („Methane Quality”) típusú szenzoroknak is nevezik és többféle változatban lettek kifejlesztve, különféle gázok érzékelésére (2.1. táblázat). Az „MQ-135” szenzor többféle olyan gáz jelenlétét képes felismerni, amik jelen vannak/lehetnek a cefében (köztük a CO₂-t és az alkoholt is), így ezekkel a szenzorokkal nem lehet biztos CO₂

¹<https://www.omega.co.uk/temperature/z/thermocouple-RTD.html>

mennyiséget mérni, bár nagyon olcsó eszközökről van szó.

Hadi és tsai. (2022) kimutatták, hogy az MQ-135 szenzor a különböző gázok esetében különböző érzékenységi válaszidővel mér, tehát (elvileg) a gázokat meg lehet különböztetni, ami ígéretesnek bizonyul további kutatásokhoz. Viszont a gázok mennyiséget mérni így sem tudjuk, tehát az erjedésnél keletkező CO₂ mennyiségének mérésére nem alkalmasak. Calvo-López és tsai. (2016) kifejlesztettek egy olcsón, de nem egyszerűen összeállítható elektrokémiai elven működő szenzort, amellyel viszonylag pontosan tudták mérni az elnyelt (nem pedig az erjedésnél képződő) CO₂ mennyiségét palackozott borban és sörben.

Pontosabb adatokat kapunk a *nem diszperzív infravörös* (NDIR) szenzorokkal, amik drágábbak az elektrokémiai szenzoroknál de kimondottan CO₂ mérésére lettek kifejlesztve. A CO₂ molekula elnyeli az infravörös fény bizonyos spektrumát (4.26 mikron), így elég azt mérni, hogy egy csövön áteresztett ilyen spektrumú infravörös fény hányad részét nyeli el a csőben lévő levegő, amiből az abban lévő CO₂ mennyiségét lehet kikövetkeztetni. Viszont, a vízmolekulák és egyes gázok (pl. a metán) is elnyelik az infravörös fényt (Biechele és tsai., 2015), ami zajossá teheti a mérést.

2.1. táblázat. MQ típusú szenzorok listája (<https://robu.in/mq-series-gas-sensor/>).

Szenzor	Érzékelt gáz(ok)	Szenzor	Érzékelt gáz(ok)
MQ-2	metán, bután, LPG, füst	MQ-131	ózon
MQ-3	alkohol, etanol, füst	MQ-135 ²	szén-dioxid (CO ₂), CO, ammónia, benzol, kén(-dioxid), alkohol, füst
MQ-4	metán, CNG	MQ-136	kén-hidrogén
MQ-5	természetes gáz, LPG	MQ-137	ammónia
MQ-6	LPG, bután gáz	MQ-138	benzol, toluol, alkohol, aceton, propán, formaldehid, hidrogén
MQ-7	szén-monoxid (CO)	MQ-214	metán, természetes gáz
MQ-8	hidrogén gáz	MQ-303	alkohol
MQ-9	CO, gyúlékony gázok		

Alkoholtartalom Alkoholtartalmat automatizáltan a tartályban, a cefre feletti térben, vagy pedig onnan az e-Orrba vezetve gázszenzorokkal tudunk mérni, pl. az MQ-3, MQ-135, MQ-138 vagy MQ-303 szenzorral (2.1. táblázat).

Alkoholtartalom mérésére alkohol-dehidrigénáz vagy alkohol-oxidáz enzimek reakcióin alapuló *bio-szenzorok*³ lettek kifejlesztve (Akin és tsai., 2010). Ezek, bár érzékenyek és viszonylag pontosak, stabilitásukat pár nap vagy pár hét használat után elveszítik. Ezért, ezeket egyenlőre még inkább csak kutatási és laboratóriumi célokra használják.

pH A cefre pH értékét, automatizáltan klasszikus üvegelektrodos szenzorral tudjuk mérni. Az interneten fellelhető, olcsó és számítógépre könnyen csatlakoztatható üvegelektro-

²<https://www.elprocus.com/mq-135-alcohol-sensor-circuit-and-working/>

³A bio-elektronikai szenzorok bizonyos anyagokra szenzitív (oxidáz vagy dehidrogenáz) enzimek reakciói, vagy azok mértéke alapján érzékelik az adott komponens jelenlétét vagy mennyiségét a mintában.

dos szenzorok⁴⁵⁶ alkoholos közegben való alkalmazásáról nem találtam információt, ezért ezek alkalmazhatósága a cefre pH értékének monitorozására további kutatásokat igényel.

Újabb, optikai-, elektrokémiai- (potenciometrikus- és voltammetrikus-) és mezőhatás tranzisztor elven működő pH szenzorok szakirodalmát és azok eredményeit tárgyalják és hasonlítják össze Vivaldi és tsai. (2021), akik arra jutottak, hogy további fejlesztések szükségesek ahhoz, hogy ezen szenzorok átvegyék az üvegelektrodos szenzorok helyét, bár speciális esetekben már beváltak (pl. Kattipparambil Rajan és tsai. (2016)). Egyrészt, a tárgyalt szenzorok többsége nagyon szűk tartományt vagy pedig inkább magasabb, 5 feletti pH értékeket képes mérni (ami nem elég a cefre pH-jának mérésére), esetleg nagyon rövid élettartamúak.

Szárazanyagtartalom A cefre glükóztartalmának mérésére alkalmasak lehetnek a különböző glükóz-oxidáz enzimeken alapuló bio-szenzorok. Érdekeseek a szabad forráskódú, viszonylag olcsón elkészíthető és számítógépre könnyen csatlakoztatható *iSpindle*⁷ és *Wifi Digital Hydrometer*⁸ megoldások, amelyek hasonló elven működnek mint ahidrométerek (fokolók): Egy súlyozott, a cefre tetején úszó hengerben elhelyezett giroszkóp méri a henger dőlésszögét, ami a hengerre ható, a cefre sűrűségétől függő felhajtóerő függvényében változik. Ezeket a megoldásokat eddig sör- és borkészítésnél tesztelték, ahol a cefre folyékony halmazállapotú. Gyümölcscefrékben illetve törkölyben való alkalmazásuk vizsgálatra szorul, valószínűleg technológiai változtatásokat is igényel.

Aroma komponensek A cefre illókomponens tartalmát e-Orr alkalmazásával vizsgálhatjuk automatizált módon. Az e-Orr (Mendez és Preedy, 2016; Dymerski, 2017), lényegében egy sor gázszenzorból áll, amelyek egyenként többféle (nemcsak egy specifikus) gázmolekulát érzékelnek, viszont együttesen képesek az adott mintához egyedi adatsorokat, ún. „ujjlenyomatokat” generálni. Ezekből az adatsorokból pedig gépi tanulás (GT) módszerekkel olyan modelleket lehet alkotni, amelyek a különböző illatokat (az adatsorokhoz rendelt címkéket) osztályozni tudják. (GT) alkalmazása az e-Orr adatokra több problémába is ütközik: Először is, kellő mennyiségű címkézett adatsorra van szükség ahhoz hogy pontos modelleket kapjunk. A másik probléma az, hogy az adatsorok inkonzisztensek lehetnek környezeti változások (pl. hőmérséklet, páratartalom) illetve a szenzorok előregedése miatt (Ye és tsai., 2021). Ez az ún. *szenzor sodródás* jelenség, amire mindenképpen oda kell figyelniünk e-Orr alkalmazásánál.

A cefre oldatban lévő aromakomponenseit e-Nyelv segítségével vizsgálhatjuk. Az e-Nyelv (Mendez és Preedy, 2016; Dymerski, 2017) ugyanazon elvek alapján működik mint

⁴<http://homelab.link/homelab-ph-rev-4.shtml>

⁵<https://atlas-scientific.com/blog/ph-sensor-raspberry-pi/>

⁶<https://www.dfrobot.com/blog-1138.html>

⁷https://github.com/universaml/iSpindel/blob/master/docs/README_en.md

⁸<https://hackaday.io/project/21386-wifi-digital-hydrometer>

az e-Orr, kivéve, hogy nem gázokat, hanem folyadékokat analizál. Itt is számolnunk kell a szenzor sodródás jelenséggel, főként a hőmérséklet változás, a szennyeződés és az ún. memória effektus miatt, amikor az előző mintából a membránon fennmaradó komponensek befolyásolhatják a mérést. Ezeket matematikai módszerekkel némiképp ki lehet küszöbölni (Kovacs és tsai., 2020), viszont, mivel a szenzoroknak érintkezniük kell a cefrével, tisztántartásuk nehezebb. Ezért, valamilyen technológiai megoldásra is szükség lenne az e-Nyelv alkalmazásánál ahhoz, hogy csak a cefre szűrt levét vezessük a szenzorokhoz. Egy másik kérdés, ami szintén vizsgálatra szorul, hogy a cefre alkoholtartalma miként befolyásolja a mérést illetve a szenzorok élettartamát. Eltérően a 2.1. táblázatban bemutatott gázszenzoroktól, nem találtam információt kereskedelembe kapható, olcsó, és kis kapacitású eszközökre csatlakoztatható e-Nyelv szenzorokról.

Egyéb szenzorok Érdekes információkat kaphatunk *optikai szenzorok*⁹ használatával, amik közül, Biechele és tsai. (2015) szerint, a *spektroszkópiás szenzorok* (pl. infravörös, UV) alkalmazása bizonyul ígéretesnek. Ezek előnye, hogy nem érintkeznek a mérendő anyaggal és egyszerre több paraméter mérésére is alkalmazhatók (Biechele és tsai., 2015; Rahmatnejad és tsai., 2023). A multi-spektrális szenzorok olcsók¹⁰ (Allepuz Rico, 2020), és kis kapacitású eszközökre is csatlakozthatók.

Adattípusok Az itt bemutatott szenzorok által mért adatokat, az egyéb kategóriában felsorolt szenzorokon kívül, *idősorok* formájában tudjuk reprezentálni. Idősornak nevezük egy változó adott intervallumonként mért értékeinek sorozatát. A spektrális kamerák komplexebb adatsorokat képeznek, amelyeket úgy képzelhetünk el, hogy egy kétdimenziós kép minden pontjához a különböző spektrumokon mért fény intenzitását rendelik, lényegében egy háromdimenziós tenzort alkotva.

2.2. Adatelemzési módszerek

Ebben a fejezetben a szenzorok által generált adatokra alkalmazható adatelemzési módszerek vannak röviden ismertetve (2.2. táblázat), alapelveikre fókuszálva (mivel csak a felhasználásukat vizsgáljuk), kihagyva a működésüknek és beállításainak leírását.

A módszerek ismertetésénél az általam oktatott, „Introduction to Data Mining” (Horváth, 2023a; Moreira és tsai., 2018) és „Machine Learning” (Horváth, 2023b; Flach, 2012; Alpaydin, 2010) MSc kurzusok előadásanyagaiból, illetve felhasznált szakirodalmából¹¹ merítettem. A különböző módszerek rendszerezése az én látásmódomat tükrözi. A szak-

⁹A *kemo-optikai szenzorok* csoportjába tartozik a fent említett nem diszperzív infravörös CO₂ szenzor.

¹⁰pl. <https://cordis.europa.eu/project/id/101028085>

¹¹Moreira és tsai. (2018) tankönyve kifejezetten nem matematikus-informatikus hallgatók részére lett írva. Flach (2012) tankönyve is jól érthető de már több matematikai tudást igényel.

irodalom, amely ebben a kérdésben nem egységes, többféle rendszerezést is használ, viszont a különböző módszereket két nagy csoportra osztja, az elemzés céljától függően.

Az egyik csoportba, amit *adatbányászat* néven is ismerünk tartoznak az ún. *deskriptív módszerek*, amik alkalmazásának a fő célja az adatokba való mélyebb betekintés, az azokban lévő összefüggések kinyerése. Itt fontos, hogy emberek által könnyen értelmezhető modellekkel dolgozzunk, hogy az ezek által kinyert és képviselt – eddig ismeretlen – információkat és tudást a későbbiekben fel tudjuk használni.

A másik csoportot, amit a *gépi tanulás* módszereinek is szoktak nevezni képviselik az ún. *prediktív módszerek*. Ezen módszerek alkalmazásának a fő célja, hogy az eddig feljegyzett megfigyeléseinkből olyan modelleket nyerjünk, amelyekkel képesek vagyunk az eddig nem ismert megfigyelések kimenetét előrejelezni. Itt már nem annyira fontos, hogy ezek a modellek emberek által könnyen értelmezhetőek legyenek, inkább a pontosságukra fókuszálunk, hogy minél jobb előrejelzésekre legyenek képesek.

2.2. táblázat. A főbb adatbányászati és gépi tanulás módszerek rendszerezése

Adatbányászat	deskriptív	statisztikai	egyváltozós	számszerű mutatók
			többváltozós	komplex mutatók
				faktor analízis
		algoritmikus	klaszterező	konvex
				nem konvex
			mintakereső	gyakori elemhalmazok
				asszociációs szabályok
Gépi tanulás	prediktív	könnyen értelmezhető	lineáris	lineáris regresszió
				logisztikus regresszió
			nem lineáris	k-legközelebbi szomszéd
				Bayes-hálók
				Markov modellek
				döntési fák
		véletlen erdők		
		gradiens fokozók		
		nehezen értelmezhető		adaboost
				szupport vektor gépek
			neurális hálózatok	

2.2.1. Adatbányászat

Az adatbányászat általános, különböző ágazatokra alkalmazható folyamatmodellje az ún. *CRISP-DM*¹² szabvánnyal írható le (Chapman és tsai., 2000), amely 6 fázisból áll, úm. i) a célok megértése, ii) az adatok megértése, iii) az adatok előkészítése, iv) modellalkotás, v) a modell kiértékelése és vi) a modell alkalmazása. Az IE alatt gyűjtött szenzor adatok

¹²A „Cross-Industry Standard Process for Data Mining” rövidítése.

elemzése szempontjából a cél az erjedés különböző paramétereinek közti összefüggések és mintázatok kinyerése az adatokból. Az adatok megértése szempontjából, azok típusait a 2.1. fejezetben ismertettük. Mivel ezek sztenderd adattípusok – amelyekkel a legtöbb modell képes dolgozni – azok előkészítése nem igényel különösebb munkát¹³. Végezetül pedig, a kapott eredmények (a modellek) kiértékelése és alkalmazása szubjektív, a főzőmestertől függ. Ezért, ebben a fejezetben inkább a modellalkotás fázisra fókuszálunk, a különböző modelleket azok komplexitása alapján rendezve.

Statisztikai Módszerek

Az adatokat általában statisztikai módszerekkel kezdjük elemezni. A cél az, hogy az adatokból olyan lényegi információkat nyerjünk ki, amelyekkel azok jól jellemezhetők.

Először az egyes változókat – úm. az egyes szenzorok által gyűjtött adatokat – vizsgáljuk külön-külön. A változók egyszerű, számszerű mutatói többek között az adatok középértékei (helyzetmutatói), szóródása, stb. Azután érdemes a komplexebb mutatókat is megvizsgálni, mint pl. a gyakorisági sorok (más néven hisztogramok). Mivel ezek már komplexebb információk, gyakran szokták ezeket grafikus formában, gráfok (pl. oszlopdiagramok) segítségével ábrázolni. Ezért ezeket a komplex mutatókat grafikus eszközöknek, míg a számszerű mutatókat analitikai eszközöknek is nevezik a statisztikában.

Az egyes változók mutatóinak vizsgálata után térünk rá ezen változók kapcsolatának vizsgálatára, két- vagy többváltozós módszerekkel. Ilyenek pl. a változók kovarianciája vagy korrelációja („normalizált” kovarianciája), esetleg a változók szóródási diagramjainak vagy közös eloszlásának vizsgálata.

A többváltozós statisztika egyik legismertebb és talán legszélesebb körben alkalmazott módszere a *faktor analízis* (FA), ami egy n változóból álló adathalmazt – azaz, egy n dimenziójú teret – egy kisebb, k dimenziójú térbe tömörít egy adott kritérium szerint optimalizálva. Ilyen kritérium lehet pl. az, hogy az eredeti adatokban lévő variancia minél nagyobb részét mentsük át a tömörített adatokba, ami alapján működik a főkomponens analízis módszere. A mátrix faktorizáció esetében pedig a cél az, hogy tömörítés alapjául szolgáló faktorok segítségével minél nagyobb pontossággal tudjuk rekonstruálni az eredeti adathalmazt. A tömörített adatok dimenziói (főkomponensek vagy faktorok) az adatsorok látens tulajdonságait összegzik. A FA sok más módszerrel kombinálható a szerint, hogy annak az eredményét az adatok előfeldolgozásához, vizualizációjához vagy predikciós modellek építéséhez használjuk.

Klaszterezés

Klaszterezés alatt azokat a módszereket értjük, amik az adatsorokat úgy rendezik különböző csoportokba (klaszterekbe), hogy az egy csoporton belüli adatsorok közti hasonlóság

¹³Feltételezzük, hogy az adatelemzést kellő adatbányászati tudással rendelkező szakember végzi.

nagyobb legyen mint a különböző csoportokba tartozó adatsorok közti hasonlóság. A klaszterező algoritmusokat több kontextus alapján osztályozhatjuk. Az eljárás alapján megkülönböztetünk i) *középpont-alapú*, ii) *hierarchikus* és a iii) *sűrűség-alapú* algoritmusokat. A klaszterekbe tartozás alapján megkülönböztetünk ún. i) *crisp* és ii) *fuzzy* algoritmusokat ahol egy adatsor csak egy klaszterbe tartozhat (crisp) vagy pedig az adatsor minden klaszterhez valamilyen 0 és 1 közti, ún. fuzzy értékkel kötődik (ahol a 0 érték azt jelenti, hogy az adott adatsor egyáltalán nem tartozik az adott klaszterbe, míg az 1 érték azt jelzi, hogy az adott adatsor 100%-ban az adott klaszterbe tartozik). A klaszterezés kimenete szerint pedig megkülönböztetünk i) *konvex* és ii) *nem konvex* klaszterező algoritmusokat, a szerint, hogy az eredményül kapott klaszterek konvexek¹⁴ vagy nem.

Minden klaszterező algoritmusnak megvannak a saját, egyedi paraméterei amiket a jó eredmény érdekében „be kell hangolni”. Egy paraméter – a legfontosabb – viszont mindig algoritmusnál megtalálható. Ez az ún. *hasonlósági függvény* ami két adatsor közti hasonlóságot adja meg. Mivel az n dimenziójú adatsorokat általában n dimenziójú Euklidészi térben reprezentáljuk, a hasonlósági függvény fordítottan arányos a pontok közti távolság függvényével, amit többféle képpen is számolhatunk (pl. Euklidészi-, Manhattan- vagy koszinusz távolság).

A legjellemzőbb szenzor adattípus az *idősor*, amely esetében a klasszikus, fent említett távolságfüggvények alkalmazása nem mindig vezet jó eredményre. Ennek az oka az adatsorok és az idősorok közti szemantikai eltérés. Itt azt kell érteni, hogy egy n dimenziójú adatsor elemei általában más-más változók értékeit tartalmazzák, míg egy n dimenziójú idősor ugyanannak a változónak n szabályos időközönként mért értékeit tartalmazza. Továbbá, a klasszikus távolságfüggvények – mivel egy adott dimenziójú térben lévő pontok (adatsorok) közti távolságot számítják – feltételezik, hogy mindegyik adatsor dimenzionalitása egyenlő. Az idősoroknál viszont ez a feltétel nem mindig teljesül, ezért a klasszikus távolságfüggvények alkalmazásához a különböző idősorokat egyforma és/vagy kisebb hosszúságú adatsorokba kell *aggregálni*, pl. az egyre több alkalmazásban használt *szimbolikus aggregált approximáció* (SAX) módszerrel, ami az idősorokat karakterorozatokká alakítja (Lin és tsai., 2007), lehetővé téve a nem idősorokra kifejlesztett (bioinformatikai vagy szövegbányászati) módszerek alkalmazását is. Az idősorok egyenlő hosszúságú vektorokra való tömörítéséhez használhatjuk a Fourier vagy wavelet *transzformációkat* is, amik az n hosszúságú idősorokat az „idő-doménből” a „frekvencia-doménbe” képezik le. Ezután, csak a legtöbb információt hordozó k frekvencia értékeit használjuk fel, hasonlóan a FA módszeréhez. Az egyik legnépszerűbb, idősorokra használt távolságfüggvény a *dinamikus idővetemítés* (DTW), amely nem azonos hosszúságú idősorokra is alkalmazható (Ratanamahatana és Keogh, 2004).

¹⁴Egy n dimenziójú térben lévő pontok halmaza akkor konvex ha a halmaz bármely két pontját összekötő egyenesen lévő pontokat is tartalmazza. Pl. a kör vagy a négyzet konvex halmazok a 2-dimenziójú Euklidészi térben, a félhold viszont már nem konvex halmaz ebben a térben.

Mintakeresés

A mintakereső algoritmusok kutatása és fejlesztése a 90-es években, az ún. *tranzakciós adatbázisok* elterjedésével gyorsult fel igazán. *Tranzakciónak* nevezzük egy adott elemhalmaz tetszőleges részhalmazát, a tranzakciók halmazát pedig tranzakciós adatbázisnak hívjuk¹⁵. A tranzakciós adatbázisokban való mintakeresés célja, hogy a különböző elemek és elemhalmazok közti összefüggéseket (mintázatokat) feltárja.

A mintázatok alaptípusa a *gyakori elemhalmazok*. Egy elemhalmazt akkor nevezünk gyakorinak ha azt a tranzakciós adatbázisban legalább k tranzakció tartalmazza¹⁶. A gyakori mintakereső algoritmusoknál a felhasználó állítja be a kívánt k „gyakorisági” értéket.

Az *asszociációs szabályok* gyakori elemhalmazok közti asszociációkat, ok-okozati összefüggéseket jellemeznek. Itt, a k „gyakorisági” értéken kívül még egy minimum b „bizalmi” értéket is meg kell adni, hogy a módszer csak a legalább b erősségű asszociációkat keresse. Egy $A \Rightarrow B$ szabályt, ahol $A \cup B$ (A és B uniója) gyakori elemhalmaz, a következőképpen értelmezzük: Ha a tranzakciós adatbázisban egy tranzakció tartalmazza az A gyakori elemhalmazt, akkor b valószínűséggel tartalmazza a B gyakori elemhalmazt is¹⁷. Itt is a felhasználó állítja be a kívánt k és b értékeket.

Ha a tranzakciók időpontját is figyelembe vesszük, akkor ún. *szekvenciákról* beszélünk. A legtöbb gyakori mintákat és asszociációs szabályokat kereső módszert a szekvenciákra is alkalmazhatjuk, megfelelően módosítva a gyakoriság definícióját.

Minél kisebb a k és b értéke annál több gyakori elemhalmazt és asszociációs szabályt találunk (és fordítva), tehát ezeket a paramétereket be kell hangolni.

Nem minden adathalmaz tranzakciós, ezért a mintakereső módszerek nem minden esetben alkalmazhatók, viszont megfelelő transzformációval a szenzor adatok előkészíthetők úgy, hogy mintázatokat is tudjunk bennük keresni (pl. Alexy és Horváth (2022) mangalicákat megfigyelő szenzorok adatain alkalmaztak mintakeresésű módszereket).

A legjellemzőbb szenzor adattípus az idősor, ahol kétféle gyakori mintázatokat különböztetünk meg, *alakzatokat* és *motvumokat*. Az alakzatok olyan rövidebb idősorok, amelyek – vagy amelyekhez nagyon hasonlóak – jól jellemeznek bizonyos (osztályba tartozó) hosszabb idősorokat (Ye és Keogh, 2009). A motívumok pedig a hosszabb idősorok azon rövidebb részei, amelyek nagyon hasonlóak egymáshoz és viszonylag gyakoriak az adott idősorban (Lin és tsai., 2002).

2.2.2. Gépi tanulás

A gépi tanulás (GT) során három bemeneti paraméterrel dolgozunk, mégpedig, i) a *tanító adathalmaz*, amely általában felcímkézett adatsorokat tartalmaz, ii) a kiválasztott

¹⁵Egy bevásárló kosár tartalma – az üzletben található termékek (elemek) részhalmaza – a tranzakció egy népszerű példája, ami az adott üzlet tranzakciós adatbázisában van tárolva.

¹⁶pl. a {kenyér} vagy a {kenyér, tej} elemhalmazokat a bevásárlókosarak legalább 10%-a tartalmazza.

¹⁷pl. ha valaki vesz vaját és tejet akkor nagy valószínűséggel vesz kenyeret is ($\{\text{vaj, tej}\} \Rightarrow \{\text{kenyér}\}$).

modell típusa, ami algoritmussal reprezentálható, és iii) egy adott *objektív függvény*, ami alapján a kiválasztott modellt, annak valamilyen tulajdonsága (pl. pontossága) szerint optimalizáljuk. A tanítás célja az, hogy kellően pontos, de ugyanakkor kellően generikus is legyen a modell, hogy az új, címkézetlen adatsorokra – a *teszt adathalmaz* adatsoraira – a valóságnak a lehető legjobban megfelelő címkéket predikálja. Ha a címkék diszkrét értékek akkor a feladat *klasszifikáció*, ha pedig valós számok akkor *regresszió*. Ma már a legtöbb modell típus képes mindkét feladatra.

Az „betanított” (optimalizált) modell az adatsorok változói és a címkék közti összefüggést definiálja. Egy modell lehet egyszerű vagy összetett, illetve emberek által könnyen vagy nehezen értelmezhető. Az, hogy melyik modellt választjuk, több tényezőtől függ és nincs olyan univerzális modell, amelyik minden adathalmazra alkalmazható lenne (lásd *nincs ingyen ebéd tétel* (Wolpert, 1996)).

Ezért, ajánlatos minden új adathalmazon több típusú modellt is kipróbálni, lehetőleg az ún. *kereszt-validálás* módszerével (Bergmeir és Benítez, 2012): Először az adathalmaz 20%-át félretesszük, ez lesz a teszt adathalmazunk. Az adathalmaz maradék 80%-át pedig felosztjuk k (általában 3, 5 vagy 10) egyenlő részre, ún. „karámokra”. Ezután, kiválasztunk egy karámot, az ún. „validáló adathalmazt”, majd a többi karámban lévő adatsorokon addig optimalizáljuk a modell típust – különböző paraméter beállításokkal – amíg a validáló adathalmazon a legjobb eredményt kapjuk. Ezt a műveletet mindegyik karámra és mindegyik modell típusra elvégezzük és kiválasztjuk azt a paraméter beállítást, amelyik – „statisztikai alapon” – a legjobb modellhez vezetett. Majd a kiválasztott paraméter beállításokkal újra tanítjuk a modell típust az összes karám adatsorait felhasználva és leteszteljük azt a teszt adathalmazon. Ezt mindegyik modell típusra, amellyel kísérletezni szeretnénk elvégezzük, és azt a modell típust választjuk ki, amelyik a legjobb eredményeket érte el. Végül, a kiválasztott modellt újratanítjuk a teljes, eredeti adathalmazon.

Lineáris modellek

Ha az adatsorok változói $(x_1, x_2, \dots, x_m)$ és a címke (y) közti összefüggés *lineáris* ($y = a_0 + a_1x_1 + a_2x_2 + \dots + a_mx_m + \epsilon$, ahol ϵ a címkézésben rejlő zajt jelöli), akkor lineáris modelleket, *lineáris regressziót* vagy *logisztikus regressziót* (ami, a neve ellenére, klasszifikációs modell) alkalmazunk. Ezek egyszerű és könnyen értelmezhető modellek, amelyeknél a paramétereik $(a_1, a_2, \dots, a_m)$ értéke kifejezi, hogy milyen erősséggel járulnak hozzá az adott változók a címke értékéhez, míg a paraméterek előjelei azt fejezik ki, hogy az adott változók pozitívan vagy pedig negatívan járulnak hozzá a címke értékéhez.

A lineáris modellek gyorsan optimalizálhatók és kis memóriaigényűek¹⁸, ezért akkor is le szoktuk ezeket futtatni amikor nem lineáris predikációs feladatokkal van dolgunk.

¹⁸Az utóbbi időben sok területen (pl. a precíziós mezőgazdaságban) nőtt az érdeklődés a kis memóriaigényű modellek iránt, amiket kis számítási kapacitású eszközökre lehet telepíteni.

Ilyenkor új, nem lineáris változókat (pl. x_1^2 , x_1x_2 , stb.) vezethetünk be az adatsorokba a modellek betanítása előtt, amik az adatokban lévő nem-lineáris összefüggéseket „kódolják”. Ebben az esetben viszont már nehezebben értelmezhetőek a modellek.

Nem lineáris modellek

A lineáris modellek – típustól függően – valamilyen komplex formában definiálhatók, lehetővé téve az adatsorok változói és a címke közti összetettebb, nem lineáris összefüggések kifejezését is. A modellek komplexitása nem feltétlenül fordítottan arányos azok értelmezhetőségével – pl. egy viszonylag terjedelmes döntési fát valószínűleg könnyebben tudunk értelmezni mint egy egyszerű szupport vektor gépet (lásd lejjebb).

Viszont, általában, minél komplexebb modellt választunk, annál több adatra van szükségünk annak betanításához. Ellenkező esetben fennál az ún. *túltanulás* veszélye, amikor a modell az adatokban lévő zajt is sikeresen magába „kódolja”, ami a tanító adathalmazon nagyon jó predikciós teljesítményhez vezet, de a teszt adathalmazon mért teljesítmény már nem lesz kielégítő. Az ún. *regularizáció* beépítése az objektív függvénybe, illetve a modell építésének folyamatába, a túltanulás veszélyét valamennyire mérsékelni tudja, viszont az elegendő adatmennyiség hiányát nem pótolja teljesen.

„Lusta” tanulás Lusta tanulásnak nevezzük azt a módszert amikor a modellt maga a tanító adathalmaz, vagy annak egy, speciálisan erre a célra kiválasztott része alkotja. A lusta tanulás legismertebb módszere a *k-legközelebbi szomszéd* alapú predikció. Itt, minden teszt adatsorhoz megkeressük az ahhoz *k* legközelebbi adatsort a tanítóhalmazból, majd ennek a *k* legközelebbi szomszédnak az átlaga (regressziós feladatnál) vagy a többségi szavazata (klasszifikációs feladatnál) adja a predikált címkét. Ahogyan a klaszterezésnél, itt is meg kell határozni egy hasonlósági függvényt. A *k*-legközelebbi szomszéd módszere, a DTW hasonlósági függvényt használva, nagyon jól alkalmazható idősorok klasszifikációjánál (Seixas Jr. és Horváth, 2020), ami fontos lehet a szenzor adatok elemzésénél.

Valószínűségi módszerek A valószínűségszámítás a GT egyik alapköve, így természetes, hogy ezen alapuló predikciós módszerek is léteznek, két fő csoportra osztva, úm. a *diszkriminatív* és *grafikus* módszerekre.

A diszkriminatív módszereknél a tanító adathalmaznál feltételezünk bizonyos eloszlást, ami nem mindig felel meg a valóságnak. Másfelől, viszont, jól kezelik az adatokban lévő „bizonytalanságokat” és esetleges változékonyságokat.

A grafikus módszerek segítségével modellezni, és vizuálisan ábrázolni is lehet az adatsorok változói közötti feltételes függőségeket. Az legismertebb grafikus módszerek a *Bayes-hálók*, amelyek a Bayes tétel alapján szerkesztik a változók és azok értékei közti függőségeket. Ezek, bár jól értelmezhetőek, általában sok számítást igényelnek, főleg sok változós

adatsorok esetén. A Bayes-hálók egyik leegyszerűsített, specifikus formája a *naív Bayes* módszer, amely feltételezi, hogy az adatsorok változói egymástól függetlenek. Ez nem mindig teljesül, viszont leegyszerűsíti és meggyorsítja a számításokat. Szekvenciális adatokon alkalmazzák jó hatékonysággal a *Markov modelleket*, amik a sztochasztikus folyamatoknál ismert Markov tulajdonságot feltételezik, úm. az adatsor jövőbeni alakulása csakis a jelen tulajdonságaitól függ, és független annak a múltban történt alakulásától.

Szabályalapú módszerek A szabályalapú módszerek olyan $A \Rightarrow B$ típusú, ok-okozati szabályokba¹⁹ tömörítik a tanuló adathalmazt, amelyek előzménye valamely (nem szükségszerűen az összes) változók értékeire vonatkozó feltételek konjunkciója, míg a szabályok következménye tartalmazza a címkéhez rendelt értéket. Itt fontos megjegyezni, hogy ezek a típusú szabályok nem azonosak az asszociációs szabályokkal, mivel a bemeneti tanuló adathalmaz nem tranzakciós adatbázis.

A szabályalapú módszerek legismertebb és legnépszerűbb típusa a *döntési fák*, amelyek a szabályokat egy fában (matematikai struktúrában) szerkesztik meg. A fának az ágaiba vannak kódolva a változók értékeire vonatkozó feltételek, a leveleiben pedig a címkék és az azokhoz rendelt értékek.

A valószínűségi modellekkel ellentétben, a döntési fáknál nincsenek az adatokkal kapcsolatban erős feltételezések. Viszont, hajlamosak a túltanulásra, amikor nagyon mély és szerteágazó fákat eredményeznek, amit a fák „metszésével” tudunk elkerülni.

Együttes tanulás A modellek egyik fontos jellemzője a *diszkriminatív képesség*, ami azt fejezi ki, hogy a különböző címkékkel rendelkező adatsorokat milyen pontosan tudják egymástól a térben elválasztani, tehát a különböző osztályok közti határokat milyen pontossággal képesek meghatározni. Az lineáris modellek diszkriminatív képessége nagyon gyenge, lényegében csak az azokat az osztályok közti határokat tudják jól meghatározni, amik síkokkal²⁰ leírhatók. Az eddig bemutatott nem lineáris modellek pedig nagy számításigényűek és hajlamosak a túltanulásra. Az *együttes tanulás* alap gondolata szerint, több egyszerű modell kombinációjával kiküszönölhetők ezek a problémák. Az együttes modelleket párhuzamosan (*bagging*) vagy iteratív módon (*boosting*) tanítjuk be.

Az egyszerű bagging módszernél a tanító adathalmazból több, egymástól kicsit eltérő mintát veszünk úgy, hogy azok mérete nagyjából az adathalmaz méretével azonos. Ezen a mintákon tanítunk be lineáris modelleket. Egy új adatsorra predikált címke pedig ezeknek a modelleknek az átlaga (regressziós feladatnál) vagy a többségi címkéje (klasszifikációs feladatnál). A *véletlen erdők* módszer a tanító adathalmazból úgy készít több

¹⁹pl. $(x_1 < 7 \ \& \ x_2 > 0.5 \ \& \ x_5 = 'a') \Rightarrow y = 1$, azaz, "ha x_1 értéke kisebb mint 7 és x_2 értéke nagyobb mint 0.5 és x_5 értéke egyenlő 'a'-val akkor az y címke értéke 1."

²⁰Az n -dimenziójú teret $n-1$ dimenziójú síkkal tudunk ketté választani – pl. egy síklapot (2-dimenziós sík) egy egyenessel (1-dimenziós sík), és egy 3-dimenziós teret egy síklappal.

eltérő mintát, hogy mindegyik csak előre meghatározott számú, véletlenszerűen kiválasztott változókat tartalmaz. Mindegyik mintán egy, előre meghatározott mélységű döntési fát tanít be. Új adatsornál pedig ezeknek a döntési fának az átlaga vagy a többségi szavazata lesz a predikált címke.

Az eredeti boosting módszer, az *adaboost* minden iterációban egy mintát készít úgy, hogy egy-egy adatsort a mintába annak „fontossága” alapján választja be (kezdésnél minden adatsor egyforma valószínűséggel lehet beválasztva a mintába). Az első iterációban betanít egy lineáris modellt, majd a tanító adathalmazban minden adatsor fontossági súlyát ezen modell hibájával arányosan módosítja. A következő iterációkban ugyanezt megismétli, tehát, ha az adott adatsoron nagyobb hibát vétett az előző modell akkor az nagyobb valószínűséggel fog bekerülni az aktuális iterációban választott mintába. Új adatsornál, ezeknek a modelleknek a súlyozott átlaga vagy a súlyozott többségi szavazata lesz az új címke, ahol a modellek súlya azoknak a tanító adathalmazon számított átlagos hibájával fordítottan arányos. A *gradiens fokozók* módszer az első iterációban egy egyszerű („alacsony”) döntési fát tanít be a tanító adathalmazon. A következő iterációkban viszont már nem az adatsorok eredeti címkéire tanítjuk be a fákat hanem az előző iterációban betanított fa hibáira. Új adatsornál, ezeknek a fának a láncolata alapján predikáljuk a címkét.

Míg a bagging módszereknél azt feltételezzük, hogy egyes modellek hibái kiküszöbölik egymást, addig a boosting módszereknél mindegyik új modell az előző modell hibáiból tanul. Ezek a modellek nagyon jó hatékonyságúak és viszonylag gyorsan és jól taníthatók abban az esetben is ha nem nagy adathalmazokkal dolgozunk.

Kernel módszerek A kernel módszerek legnépszerűbb képviselői a *szupport vektor gépek*, amelyeket – bár lineáris modellekről van szó – olyan feladatoknál használunk ahol a változók és a címkék közti összefüggés nem lineáris. Erre *kernel* függvények segítségével képesek, amelyek az n -dimenziójú adatsorokat egy olyan, magasabb dimenziójú térbe képezik le, amelyben a változók és a címkék közti összefüggést már lineárisan lehet modellezni. A szupport vektor gépek kis számításigényű módszerek és viszonylag kevés paraméterük van, amik beállításra és finomhangolásra szorulnak. Ennek ellenére nagyon jó diszkriminatív képességgel rendelkeznek. A modellek is kis memóriaigényűek, így kis kapacitású eszközökön is telepíthetők.

Neurális hálózatok A neurális hálózatok fejlesztését az élőlények agyi működésének számítógépes szimulációja motiválja, amiben a neuronokat több bemenettel és egy kimenettel rendelkező számítási egységeként definiáljuk, és amelyek a bemenetek lineáris kombinációját egy ún. *aktivációs függvény* útján kimenetté transzformálják. A neuronok rétegekbe rendezett hálózatot alkotnak. A neurális hálózatok egy, az adatsorok változóinak megfelelő bemeneti réteggel és egy, a címkéknek megfelelő kimeneti réteggel, illetve

néhány közbülső, ún. rejtett réteggel rendelkeznek. Egy neurális hálózat, lényegében, egyszerű matematikai függvények láncolata, amely segítségével – Hornik és tsai. (1989) univerzális közelítési tétele alapján – bármilyen komplex függvényt implementálhatunk.

A neurális hálózatok az utóbbi évtizedben kerültek igazán a figyelem középpontjába, az ún. *mély neurális hálózatok* sikereinek köszönhetően. Szinte minden adattípusra és feladatra vannak már kifejlesztett neurális hálózatok. Képeknél *konvolúciós neurális hálózatokat*, szöveges adatoknál és idősoroknál *rekurrens neurális hálózatokat*, adat generáláshoz pedig *generatív párharc hálózatokat* használhatjuk.

A mély neurális hálózatokban nem ritkán több száz millió neuron van több rétegbe rendezve. Ezek a modellek nagyon jó teljesítményre képesek, viszont a betanításukhoz hatalmas adatmennyiségre és nagyon nagy számítási kapacitásra van szükség²¹. Kis adatmennyiség, illetve kis számítási kapacitás esetén megoldást jelenthet az ún. *transzfer tanulás* alkalmazása, aminél egy meglévő, a mi adatainkhoz hasonló adatokon betanított mély neurális hálózat utolsó pár rétegét finomhangoljuk a mi adatainkat használva. Sajnos, még az ilyen transzfer tanulással finomhangolt hálózatok is hatalmasak, így azokat nem tudjuk kis memóriájú eszközökre telepíteni, valamint a kis adathalmazokon betanított mély neurális hálózatok hajlamosak a túltanulásra. Ezért, jól meg kell fontolni, hogy milyen feladathoz alkalmazzuk ezeket a módszereket.

2.3. „Survey” cikkek

2.3.1. Erjesztés monitorozása a bio- és gyógyszeriparban

A survey publikációk többsége a biogyógyszerészeti folyamatok monitorozásának szakirodalmát elemzi (mint pl. Biechele és tsai. (2015) vagy Rahmatnejad és tsai. (2023)), pár olyan paramétert is beleértve amiket az erjesztésnél is monitorozunk (pl. a pH vagy a CO₂), viszont inkább azokra a paraméterekre fókuszálva (mint pl. a tenyésztett sejtek állapota) amiket a szeszgyártásban nem szoktunk monitorozni – bár ezek is érdekesek lehetnek kutatás-fejlesztési oldalról (pl. gyümölcs-specifikus élesztő kitenyésztésénél). Ezekben a survey cikkekben összesen csak két olyan idézetet találtam amiben szeszszesitaloknál is alkalmazták a kifejlesztett szenzorokat alkoholtartalom vagy CO₂ mérésére, de ezek is inkább a szenzorok biológiai (Akin és tsai., 2010) vagy kémiai (Calvo-López és tsai., 2016) elveit tárgyalják.

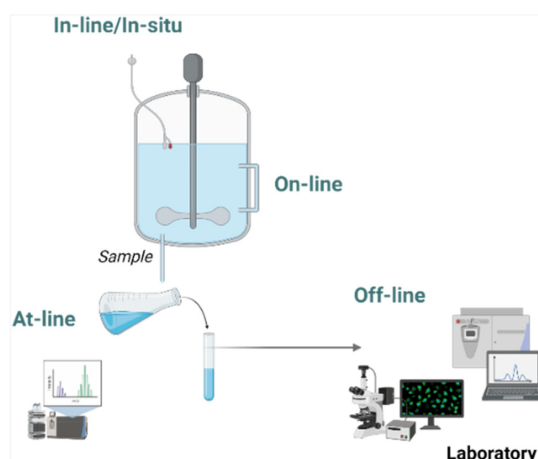
A dolgozat szempontjából érdekesebbek azok a survey cikkek amik az általános erjesztés szenzorokkal való monitorozásának szakirodalmát elemzik, főként bio- és gyógyszeripari szemszögből.

Cargalo és tsai. (2022) átfogó képet nyújtanak, általánosan az IE monitorozásának

²¹Ezért, ezeket a modelleket főleg a nagy tőkével és sok adattal rendelkező cégek képesek csak fejleszteni, ami egyre nagyobb aggodalmat vált ki a kutató közösségekben és a társadalomban egyaránt.

különböző módjairól, azokat négy csoportra osztva (2.1. ábra): az *in-line* (vagy *in-situ*) monitorozásnál az erjesztő tartályban vannak elhelyezve a szenzorok, a cefrében (pl. hőmérséklet, pH) vagy az erjesztőtérben (pl. CO_2). Az *on-line* monitorozásnál egy „kitérőben” vannak elhelyezve a szenzorok, amik szintén automatizált mérést végeznek (pl. pH, szárazanyagtartalom, alkoholtartalom, aromakomponensek), viszont hosszabb reakcióidővel mint az *in-line* monitorozás esetében. Ennek a módszernek az előnye, hogy gyors, a benne lévő szenzorok (pl. e-Nyelv) könnyen sterilizálhatók, viszont az erjesztőtartályt módosítani kell a mintavevéshöz és méréshez szükséges kitérővel (a mintát vagy vissza engedjük a tartályba vagy pedig kidobjuk). Az *at-line* monitorozásnál periodikusan mintát veszünk a tartályból és azt a helyben lévő műszerekkel analizáljuk (pl. savtartalom). Az *off-line* monitorozásnál pedig a levett mintákat laboratóriumban analizáljuk (pl. baktériumok, élesztők jelenléte). Minél összetettebb monitorozási módot alkalmazunk (az itt bemutatott sorrendben), annál hosszabb válaszidővel, de annál komplexebb vizsgálatokat végezhetünk.

2.1. ábra. Az IE monitorozásának módjai (forrás: Cargalo és tsai. (2022)).



2.3.2. e-Orr és e-Nyelv alkalmazása az élelmiszeriparban

Az egyre elérhetőbb technológiai megoldások lehetővé tették az e-Orr és e-Nyelv széleskörű alkalmazását az élelmiszeriparban. Ezek a survey cikkek már több szeszgyártással kapcsolatos cikket idéznek mint a fent említett, bio- és gyógyszeripari alkalmazásokra összpontosító survey cikkek. A 2.3. táblázat tartalmazza azon idézetek számát, amelyek ezekben a survey cikkekben vannak elemezve és szeszgyártással, illetve a dolgozat témájával, az alkoholos erjesztéssel kapcsolatosak.

Tan és Xu (2020) szerint az e-Orr használata nem kontrollált mintavételnél nehéz, mivel az illókomponensek mennyisége hőmérséklet-, páratartalom- és nyomásfüggő. Nehezíti még az alkalmazást az, hogy nagy mennyiségű czmkézett adatra van szükség ahhoz, hogy jó GT modelleket taníthassunk be, amin sokat segítene egy szabadon elérhető információ

tár. Azt is nehezményezik, hogy a szakirodalomban fellelhető mérések nehezen összehasonlíthatók, mert sok különböző típusú érzékelőt²² tartalmazó szenzort, különböző minta előkészítési eljárást, tanító adathalmazt és GT módszert használnak. Ezért ajánlják a standardizált szenzorokkal való mérési adatok gyűjtését. Ezeket a hiányosságokat az e-Nyelv esetében is említik, hozzátéve az ún. „mátrix effektust”, amikor a valódi mérések eredményei eltérnek a laborban, steril mintákon mért mérések eredményeitől.

Átfogó, makroszkopikus képet nyújt az utóbbi három év e-Orr élelmiszeripari alkalmazásairól Wang és Chen (2024), akik arra hívják fel a figyelmet, hogy jobban oda kellene figyelni az e-Orr kiválasztásánál, és (lehetőség szerint) az adott feladatnak megfelelő berendezést kell választani vagy építeni – a barkácsolt szenzorokat flexibilisebben tudjuk összeállítani, viszont alapos tesztelésnek kell ezeket alávetni. Arra is figyelmeztetnek, hogy a legtöbb kutatás nem veszi figyelembe a zajt és a szenzor sodródást (2.1. fejezet), így nem is használnak megfelelő korrekciókat, ami miatt a kapott értékek és eredmények pontatlanok lehetnek. Végül, megkérdőjelezi a modern GT módszerek jelenlegi, nagyon kis mintaszámú mérési adatokon való alkalmazásának szükségességét, mivel ezek a módszerek nagyon nagy adatmennyiséget igényelnek a betanításukhoz, amely begyűjtéséhez még időre és sztenderdizált folyamatokra van szükség.

Az e-Orr alkalmazásait tárgyalva, Karakaya és tsai. (2020) röviden kitérnek a szeszipari alkalmazásokra is, és felkeltik a figyelmet a *bioelektronikus* orrokra (Lim és Park, 2014), amik biomolekulák és nanotechnológia segítségével érzékelik az illóanyagokat, viszont ezen szenzorok elterjedésére még várunk kell.

Cheng és tsai. (2021) nagyon részletes elemzést nyújtanak a *kompakt e-Orr* fejlesztésének és kutatásának területéről és jövőbeni lehetséges irányairól. 31 survey cikket is elemeznek kronológiai sorrendben, amelyeknél az elmúlt évtizedben a hangsúly inkább az e-Orr alkalmazásaira tolódott ki. A fém-oxid félvezető típusú szenzorok (lásd. 2.1. táblázat) a legtöbbet használt kompakt e-Orr szenzorok kifejlett technológiájuk, egyszerűségük, széles spektrumuk és kedvező árak miatt. A kompakt e-Orr kis mérete és alacsony energiafelhasználása miatt érdekes a dolgozat témájának szempontjából, mivel az egyszerűbben alkalmazható több, különböző típusú és méretű erjesztő tartály esetében (ami általános egy átlagos főzdeben). Wang és Chen (2024) véleményéhez hasonlóan, a klasszikus GT módszerei mellett is érvelnek, nem feltétlenül a rendelkezésre álló kis adatmennyiség miatt, hanem azért mert ezek a módszerek kis eszközökre is telepíthetők, megkönnyítve az *on-line adatfeldolgozást*, ahol a számításokat vagy azok egy részét magán az eszközön található processzorok végzik. Ez lehetőséget ad arra, hogy a különböző érzékelők a *dolgok internetébe*²³ legyenek bekötve és kommunikálni tudjanak egymással.

A dolgok internete erjesztési folyamatokban való alkalmazásait tárgyalják Adeleke és tsai. (2023), 26 kiválasztott publikációt elemezve, amik közül 12 érinti a szeszgyártást.

²²pl. polimerek, fém-oxid félvezetők, kvarckristály mikromérlegek, felület-akusztikus hullámérzékelők.

²³Egymással kommunikáló, egyszerű, kis számítási kapacitású elektronikai eszközök rendszere/hálózata.

Ezek közül 3 cikkben a baiju²⁴ erjesztését monitorozzák, amik a 2.4. fejezetben lesznek elemezve. A baiju erjesztése – a többi szeszessital erjesztésétől eltérően – szilárd fázisú fermentációval²⁵ történik, ahol olyan paramétereket is monitorozni kell (pl. páratartalom), amiket a többi szeszessital esetében nem szükséges.

Ye és tsai. (2021) és Liu és tsai. (2023) a szenzor adatok elemzésének kérdéseit tárgyalják, különös figyelmet szentelve a szenzor sodródás és az adatokban lévő zaj kezelésének. Ye és tsai. (2021) megjegyzi, hogy a kevés adatmennyiség ellenére, a transzfer tanulást felhasználva, a modern GT módszerei jó eredményeket értek el, viszont egységes, kellően nagy mennyiségű és szabadon hozzáférhető tanító adatok híján nehéz a szakirodalomban közzétett eredményeket összehasonlítani, mivel ezen kutatások eredményei saját adatokon történtek. Azt a kérdést is felvetik a transzfer tanulás elveire hivatkozva, hogy a különböző gázoknál mért mintázatok az adatokban átvihetők lehetnek olyan gázok elemzéséhez, amikről kevés adat áll rendelkezésre. Liu és tsai. (2023) szerint a jelenlegi adatelemzési módszereknél – mivel azok főként laboratóriumi körülmények közt gyűjtött adatok alapján lettek kifejlesztve – jobban figyelembe kellene venni a valós körülmények közti adatgyűjtésnél fellépő zavaró hatásokat, mint pl. más aromakomponensek vagy idegen szagok interferenciáját. Továbbá, az interpretálható modellek alkalmazásának előnyeit is kiemelik, amik segítségével mélyebb betekintést nyerhetünk az adatokba, ami elősegítheti a további fejlesztéseket. Végül pedig, Wang és Chen (2024) és Cheng és tsai. (2021) megjegyzéseikhez hasonlóan, az olyan GT módszerek alkalmazását is javasolják (a modern, mély tanulás módszereinek alkalmazása mellett), amelyeket a kompakt e-Orr-okba is integrálni lehet.

Cargalo és tsai. (2022) tárgyalják még az ún. *szoft szenzorok* jelentőségét is. Szoft szenzoroknak nevezzük azokat a szoftveres megoldásokat, amelyek segítségével az erjesztés egyszerűen és gyorsan mérhető paramétereiből becsüljük meg annak nehezen mérhető paramétereit, a folyamatokat modellező matematikai modellekkel vagy GT módszereivel Zhu és tsai. (2020). A MI adta lehetőségeket az erjesztés monitorozásában és irányításában (aminek része a GT is), egyébként már nagyon régen felismerték (Montague és tsai., 1989). A szoft szenzorok kifejlesztésénél figyelembe kell venni, hogy az erjesztés különböző matematikai- modelljeit többnyire kontrollált, laboratóriumi környezetben mért adatok alapján dolgozták ki, így azok, valószínűleg kiigazításra szorulnak.

Egy régebbi cikkben Baldwin és tsai. (2011) megjegyzi, az e-Orr és e-Nyelv alkalmazásait elemezve sör és bor klasszifikációjánál, hogy az alkohol és a CO₂ jelenléte megnehezíti az e-Orr alkalmazását. Továbbá, ajánlják ezen szenzorok kombinálását más szenzorokkal – pl. fotodetektorral (*elektronikus szemmel*) a jobb eredmény elérése érdekében.

²⁴A baiju a világ egyik legnépszerűbb, a távol-keletről származó szeszessitala. Főként cirokból készül de vannak rizsből vagy kölesből (esetleg búzából, árpából) készült változatai is.

²⁵Szilárd-fázisú fermentációnál nem folyékony közegben történik az erjesztés (pl. kenyér kelesztése).

Peris és Gilabert (2013) megjegyzik, az e-Orr, de főleg az e-Nyelv alkalmazásai még a kezdeti stádiumban vannak és az ipari méretű elterjedésükhöz további, rövid és hosszútávú kutatások szükségesek, sokkal magasabb mintaszámokkal és szenzor típusokkal. Megjegyzik továbbá, hogy az e-Orr-al inkább kvalitatív, míg az e-Nyelv kvantitatív elemzésekhez is használható.

2.3.3. e-Orr és e-Nyelv alkalmazása a szesziparban

Az e-Orr különböző alkalmazásait a sörgyártásban tárgyalják röviden Santos és tsai. (2017), Varnamkhasti és tsai. (2011). Santos és tsai. (2017) megállapítja, hogy a legtöbb alkalmazás az alapanyagok és a késztermék vizsgálatával, valamint a különböző sörök megkülönböztetésével kapcsolatos. Varnamkhasti és tsai. (2011) az etanolnak és víznek az e-Orr működését zavaró hatását elemzik részletesen. Az etanolra több szenzor is érzékeny, ami befolyásolhatja a mérést, vagy elnyomhatja a kisebb koncentrációban lévő komponensek felismerését, amit megfelelő referencia mintákkal és kalibrációval lehet kiküszöbölni. A víz zavaró hatását szilárd fázisú mikroextrakciós mintavétel²⁶ alkalmazásával javasolják kiküszöbölni, aminek a teljes automatizálása egyenlőre költséges megoldásnak bizonyul.

A e-Orr és e-Nyelv élelmiszeripari alkalmazásait foglalják össze Mendez és Preedy (2016), több fejezetet is szentelve kimondottan a szeszipari alkalmazásoknak: A különböző fő ízeket (édes, sós, savanyú, keserű, fanyar, csípős és az „umami”) tárgyalják a 16.-ik fejezetben Toko és Tahara (2016), olyan e-Nyelv számára fejlesztett lipid/polimer membránokat (és azok laboratóriumban való elkészítésének módját) bemutatva, amelyek képesek ezeket az ízeket érzékelni²⁷. A 14. fejezetben Lozano és tsai. (2016) az e-Orr borászati alkalmazásait tárgyalják, amit a vízen és az etanolon kívül még a kén²⁸ jelenléte is megnehezít. Továbbá, bemutatnak egy általuk kifejlesztett mintavételi eljárást, aminek segítségével folyamatosan lehet monitorozni a bor állapotát az érlelő tartályban, amihez a tartályt kissé módosítani kell. Külön részben tárgyalják a bor hibáinak érzékelésével kapcsolatos cikkeket, mint pl. az ecetsav (Macías és tsai., 2013b) vagy gombák okozta hibák (Cabanes és tsai., 2009) érzékelésére tett kísérleteket. Lozano és tsai. (2016) két cikket is idéznek, amik érdekesek lehetnek az etanol és víz okozta nehézségek kiküszöbölésére. Az egyikben gázkromatográffal választják el a vizet és az etanolt, és csak a többi komponenst engedik az e-Orr-ba (Ragazzo-Sanchez és tsai., 2009), míg a másikban egy nedves²⁹ e-Orr” segítségével méri a bor savtartalmát (Gil-Sánchez és tsai., 2011). A 26. fejezetben Rodríguez Méndez és tsai. (2016) a borkészítés egyes fázisaira lebontva tárgyalják az e-Nyelv

²⁶Vízre és etanolra alacsony affinitású, nedvszívó gyanták segítségével gyűjtjük össze a minta fejterében lévő illóanyagokat (https://en.wikipedia.org/wiki/Solid-phase_microextraction).

²⁷Érdemes lenne kipróbálni az édességet érzékelő szenzort (Toyota és tsai., 2011) a cefre cukortartalmának monitorozására, még ha egyenlőre csak laboratóriumi környezetben is.

²⁸Egyes gyümölcsök (pl. a meggy) cefréjének kéntartalma megnehezítheti a mérést.

²⁹Az elektródák egy nedves textilben helyezkednek el, ami magába szívja az illókomponenseket.

alkalmazásait, külön kitérve az erjesztés monitorozására, és megjegyzik, hogy az e-Nyelv alkalmazása nehézségekbe ütközhet az erjedés turbulens volta miatt, így azt más szenzorokkal ajánlják kombinálni (Buratti és tsai., 2011). A bor organoleptikus tulajdonságaira kitérve, idézik Legin és tsai. (2003) cikkét, akik e-Nyelv segítségével, viszonylag pontosan tudták mérni a bor organoleptikus paramétereit, többek közt a pH-t és az etanoltartalmat. A 29. fejezetben Branchini és tsai. (2016) az e-Orr és az e-Nyelv együttes alkalmazásainak előnyeit méltatják, lehetőség további, szerint más típusú szenzorokkal (pl. elektronikus szemmel) kiegészítve. A különböző érzékelők által gyűjtött, komplementáris információk fúziójával³⁰ hatásosabb érzékelést érhetünk el (Rodríguez-Mendez és tsai., 2004).

Az e-Orr és az e-Nyelv borászati alkalmazásait tárgyalják Rodríguez-Méndez és tsai. (2016), a holisztikus megközelítés mellett érvelve, amikor különböző elv alapján működő e-Orr-okat és e-Nyelv-eket kombinálunk a minél hatékonyabb érzékelés eléréséhez. Azt is megjegyzik, hogy nagy szükség lenne egy közös nemzetközi, és szabadon hozzáférhető adatbázis létrehozására a különböző szenzorokkal végzett kísérletek adataival, amely nagyban elősegítené a további kutatást és fejlesztést ezen a területen. Javasolják a *bioszenzorok* jövőbeni minél szélesebb körű alkalmazását, aminél viszont arra hívják fel a figyelmet, hogy minden enzim adott körülmények közt funkcionál, ezért nem mindegy, hogy azokat milyen módon kombináljuk.

Vasilescu és tsai. (2019) részletesen tárgyalják a bio-szenzorok borászatban való alkalmazásainak szakirodalmát és az abban bemutatott bioszenzorok kémiai hátterét, különös figyelmet fordítva a bor glükóz- és fenoltartalmának megállapítására, és elemzik a piaci bioszenzor-alapú monitorozó rendszereket, amelyek (egyenlőre) drágák és körülményes a használatuk. Bemutatnak egy általuk kifejlesztett monitorozó rendszert kis- és közepes borászatok részére (ennek az ára megközelítőleg 15e Euró).

Az e-Orr alkalmazását az élelmiszeripari erjesztés monitorozására tárgyalják Seesaard és Wongchoosuk (2022), külön fejezetet szentelve az alkoholos erjesztésnek, a tejsavas (pl. jogurt, szalámi vagy savanyú káposzta), ecetsavas (pl. ecet, kávé, tea) és a lúgos (pl. miso, hal- vagy szója szósza) erjesztés mellett. Az alkoholos erjesztés monitorozásánál tárgyalják a kenyér kelesztésének monitorozását is, ahol különböző, a cefre erjedésénél is keletkező komponensek (mint pl. az acetaldehid, metanol vagy a különböző, magasabb rendű alkoholok) jelenlétéből állapítható meg a termék frissessége. Megjegyzik, hogy jelenleg az e-Orr főként a termékek minőségének és hitelességének megállapítására van alkalmazva de idéznek egy pár, az erjesztés monitorozásával foglalkozó cikket is. Megállapítják, hogy a legtöbb kutatás fókuszában az e-Orr méretének, mintavételi módszerének, illetve az adatfeldolgozás algoritmusainak fejlesztése áll (ez utóbbinál, szintén érintik egy nemzetközi adatbázis létrehozásának szükségességét). E helyett, inkább az újabb és érzékenyebb gáz szenzorok fejlesztésére kellene koncentrálnia a jövőbeni kutatásoknak.

³⁰Az ún. szenzor fúzióval az emberi érzékelést imitáljuk, amikor az italt előbb a szemünkkel, majd az orrunkkal és a szájunkkal érzékeljük, így jobb képet alkotva annak minőségéről.

Mendez és Preedy (2016) 28.-ik fejezetében Buratti és Benedetti (2016) az e-Orr és az e-Nyelv kimondottan élelmiszeripari alkoholos erjesztésre való alkalmazásait tárgyalják, és bemutatnak egy esettanulmányt amiben a must paramétereinek erjedés közbeni változásait voltak képesek monitorozni a két szenzor fúziójával és egyszerű statisztikai modellek segítségével, viszont megjegyzik, hogy több kutatásra van még szükség.

Végezetül, Violino és tsai. (2020) az alacsony költségű és (lehetőség szerint) szabad forráskódú és szabad hardverű technológiákat tekintik át, amelyek a sörfőzés teljes láncát lefedik, és amiket a „sör internetének” hálózatába lehet csatlakoztatni. Megjegyzik, hogy bár ezek a technológiák ígéretesek, a kellő szakirodalom hiányában tudományosan még nem bizonyítottak az előnyeik. Főleg internetes forrásokat idéznek. Részletesen tárgyalják a nyomonkövethetőség és átláthatóság³¹ kérdését is, amit a *blokklánc*³² technológia fejlődése tesz lehetővé.

2.3. táblázat. Survey cikkek szenzorok élelmiszeripari alkalmazásairól a bennük található összes, a szeszesitalokkal kapcsolatos és az alkoholos erjesztéssel kapcsolatos idézett cikkek számával („eO” = e-Orr, „eN” = e-Nyelv, „eSz” = egyéb szenzor).

Survey	Szenzor			Idézetek száma		
	eO	eN	eSz	összesen	szeszesital	erjesztés
Cargalo és tsai. (2022)			+	266	5	1
Adeleke és tsai. (2023)			+	52	12	5
Tan és Xu (2020)	+	+		106	5	–
Wang és Chen (2024)	+			368	8	–
Karakaya és tsai. (2020)	+			297	8	–
Baldwin és tsai. (2011)	+	+		110	13	–
Cheng és tsai. (2021)	+			210	13 ³³	–
Ye és tsai. (2021)	+			125	9 ³³	–
Liu és tsai. (2023)	+			297	1	–
Santos és tsai. (2017)	+			77	35	2
Varnamkhasti és tsai. (2011)	+			73	27	–
Mendez és Preedy (2016)	+	+	+	220 ³⁴	115 ³⁴	3
Rodríguez-Méndez és tsai. (2016)	+	+		105	73	4
Vasilescu és tsai. (2019)	+	+		120	67	12 ³⁵
Seesaard és Wongchoosuk (2022)	+			198	24	2
Buratti és Benedetti (2016)	+	+		45	17	12
Peris és Gilabert (2013)	+	+		42	5	1
Violino és tsai. (2020)			+	–	–	–

³¹A mezőgazdaságban és élelmiszeriparban jól ismert „farmtól az asztalig” koncepciót követve, ami lassan be fog gyűrűzni a szeszgyártás területére is.

³²<https://hu.wikipedia.org/wiki/Blokklánc>

³³Pár cikkben a bor ecetesedését vizsgálják (cefre ecetesedésének detektálásához hasznos lehet).

³⁴A könyv szeszgyártást érintő négy fejezetében (Toko és Tahara, 2016; Lozano és tsai., 2016; Rodríguez Méndez és tsai., 2016; Branchini és tsai., 2016) lévő összes idézet száma (lehetnek átfedések).

³⁵Ezekből 6 cikk malolaktikus fermentáció monitorozásáról szól.

A 2.3. táblázatban a survey cikkek vannak összefoglalva, feltüntetve az ezek által idézett összes cikk számát, a szeszesitalokkal, illetve a szeszesital erjesztéssel kapcsolatos cikkek számát³⁶ – beleértve pl. a malilaktikus erjesztést vagy olyan cikkeket, amik pl. az erjedés vagy a hibás erjedés melléktermékeinek (pl. ecetsav, acetaldehid) érzékelésével foglalkoznak. A táblázat 6. oszlopából kitűnik, hogy kevés cikk érinti a szeszipari alkalmazásokat (az összes cikk kb. 16%-a), amik leginkább két csoportba sorolhatók:

Az első csoportban vannak klasszifikációs feladatokat érintő cikkek, amik az italokat osztályozzák különböző csoportokba aromaprofilok alapján, pl. a sör típusának predikciója vagy a bor eredetiségének megállapítása (pl. Aleixandre és tsai. (2008), Blanco és tsai. (2015), Chen és tsai. (2023), Gutiérrez és tsai. (2013), Li és Gu (2023), Qi és tsai. (2015), Qi és tsai. (2017b), Qi és tsai. (2017c), Qi és tsai. (2017a), Santos és tsai. (2011), Wang és tsai. (2021), Yang és tsai. (2020), Zhao és tsai. (2017), Zhang és tsai. (2021) és Zhang és tsai. (2022)).

A másik csoportba tartoznak azok a cikkek, amelyek az italok minőségét vagy valamilyen tulajdonságait mérik, pl. a bor cukortartalmát vagy a sör fertőzőittségét (pl. Ghasemi-Varnamkhasi és tsai. (2012), Gonzalez Viejo és tsai. (2020), He és tsai. (2021), Lorwongtragool és tsai. (2011), Liu és tsai. (2019), Macías és tsai. (2013a), Nery és Kubota (2016), Paknahad és tsai. (2017), Shi és tsai. (2019), Shi és tsai. (2008) és Wang és tsai. (2022)).

2.4. Alkoholos erjesztés in-line és on-line monitorozása

A 2.3. táblázat utolsó oszlopából látszik, hogy kimondottan szeszesital erjesztés monitorozásával nagyon kevés kutatás foglalkozik (a szesziparral kapcsolatos cikkek 13%-a, ami az összes cikk kb. 2%-a). Ezen kutatások egy része at-line vagy off-line méréseken alapul (2.1. kép), nem automatizálható laboratóriumi eszközök igénybevételével (pl. gázkromatográf vagy tömegspektroszkóp). Ezért, ebben a fejezetben csak az alkoholos erjesztés on-line és in-line monitorozásával foglalkozó kutatások kerülnek elemzésre (2.4. táblázat).

2.4.1. Szín és fenoltartalom

Shrake és tsai. (2014) bemutatták, hogy az általuk készített LED színmérő (6 látható és 2 UV fénysávban), vörösborok színét és fenoltartalmát mérve, az UV-spektrofotometriás mérésekhez hasonló eredményeket ad. A mérésekhez egy 2 μm -es membránszűrő segítségével vonták ki a mintákból az élesztősejteket és a CO_2 -t.

Tomtsis és tsai. (2016) „okos hordó” rendszerében LED-es RGB színérzékelőt használtak a bor színének és tisztaságának mérésére erjedés közben.

³⁶Pl. az első sorban, Cargalo és tsai. (2022) összesen 266 cikket idéz, ebből 5 cikk szeszesitalokkal kapcsolatos, és ebből az idézett 5 cikkből 3 érinti az alkoholos erjesztés monitorozását.

Albanese és tsai. (2011) teljesen automatizált, átfolyó rendszerű³⁷ mérőműszert fejlesztettek ki, amelyben torna-peroxidáz alapú bioszenzorral (e-Nyelv) mérték – a gázkromatográfiai vizsgálatokhoz viszonyítva nagyon pontosan – a must fenoltartalmát. A mustot hígították, hogy az e-Nyelv-ben használt szenzoroknak a mért komponensekre adott válaszreakcióik lineárisak legyenek.

Vasilescu és tsai. (2019) egy előzőleg kalibrált és korrigált szénelektroda segítségével mérték a kékszőlők mustjának fenoltartalmát, amely erős korrelációt mutatott a must színével.

2.4.2. Szárazanyagtartalom

Becker és tsai. (2001) egy 300m³-es sör erjesztő tartályban mérték a cefre sűrűségét, a tartályon kívülre szerelt, ultrahangot kibocsátó és mérő piezo kristály szenzorral. A szenzort a leülepedett élesztő réteg (a tartály kónuszos része) felett helyezték el, ahol már nem olyan sűrű a cefre viszont még nem olyan sok a CO₂, amik torzíthatják a mérést. A méréseket a cefre hőmérséklete alapján kalibrálták, a módszert később továbbfejlesztették (Hoche és tsai., 2016)³⁸.

Albanese és tsai. (2011) rendszerében glükóz-oxidáz alapú bioszenzort (e-Nyelv) fejlesztettek ki a must cukortartalmának pontos mérésére.

A Vasilescu és tsai. (2019) által kifejlesztett in-line monitorozó rendszer az erjedő mustban, az általuk készített bioszenzorok segítségével méri a glükóztartalmat (glükóz-oxidáz alapú), az összes redukáló cukortartalmat (arany elektródás katalitikus szenzor), viszonylag pontos mérési eredményeket produkálva az alkalmazott enzimes kolorimetriás referenciamódszerhez képest. A rendszer része egy automatikus hígító modul, amely a mintákat az erjedés menetének megfelelően, 1:1000 (az elején) és 1:50 (a végén) arányban hígítja, a szenzor válaszreakcióinak linearitása miatt. A szenzorokat 5 naponként cserélték.

Lidén és tsai. (2000) az erjesztőtérből kivezetett gázokban, e-Orr által mért adatokon és a cefrében, folyadékkromatográf által mért címkéken betanított neurális hálózatok segítségével predikálták, viszonylag pontosan a cefrében mért glükóztartalmat. A cefre glicerintartalmát is predikálni tudták, de már kisebb pontossággal. A kísérleteket szintetikus közegen (cefrén) végezték.

Lidgren és tsai. (2006) az általuk kifejlesztett és szabadalmaztatott SIRE[®] technológiával mérték a szintetikus cefre glükóztartalmát. A SIRE[®] bioszenzor lényege, hogy az enzimek nincsenek kimerevítve az érzékelő elektródákra, hanem a rendszer külön adagolja azokat egy membránon keresztül. Így a szenzor, különféle enzimekkel feltöltve flexibilisebben használható.

³⁷Flow Injection Analysis: a minta hordozó oldatba van fecskendezve és az abban lévő reagens anyaggal elegyítve kerül a detektorba.

³⁸Csak a cikk kivonatához tudtam hozzáférni.

2.4.3. Szén-dioxid

Buonocore és tsai. (2021) az erjesztő tartályok alá helyezett elektronikus mérleg segítségével mérték a tartályból kivezetett CO₂ súlyát (90 liternyi sör cefréjének a súlya az erjesztés végére kb. 3–3.5 kg-al lett kevesebb).

Cañete-Carmona és tsai. (2020) egy speciális alagútba helyezett infravörös CO₂ szenzorral monitorozták a bor erjesztésénél keletkező CO₂ mennyiségét.

Tomtsis és tsai. (2016) „okos hordó” rendszerében nyomásmérővel mérik az erjesztőtérben lévő CO₂ nyomását. A kotyogón áthaladó buborékok számát egy optikai szenzorral (fotodiódával) érzékelik.

Lidén és tsai. (2000) infravörös optikai szenzorral mérték az erjesztőtérből kivezetett gázokban a CO₂ tartalmat.

Mamolar-Domenech és tsai. (2023) egy (szintetikus) cefrébe eresztett hidrofonnal mérték az erjedés által keletkező CO₂ buborékok által kibocsátott hangokat az 1-4.5 kHz tartományban, amelyek a buborékok nagyságával korrelálnak³⁹. A kapott adatokon jól követhető volt az erjedés menete.

Egy nagyon fontos paramétert, a sörfőzde légterében lévő CO₂ mennyiségét mérték Arduino⁴⁰, egyszerű CO₂ szenzorok és vezeték nélküli technológia használatával Hawchar és tsai. (2022). Felhívják a figyelmet az alkalmazott vezeték nélküli technológiák áramellátására⁴¹, a bemutatott technológia 8-10 naponkénti elemcserére szorult.

2.4.4. Hőmérséklet

Tamo és Hilario-Tacuri (2020) egyszerű, Raspberry PI⁴² által vezérelt hőmérséklet szenzorokkal mérték a sör cefréjének hőmérsékletét.

Tomtsis és tsai. (2016) „okos hordó” rendszerében két hőmérőt is használnak, az egyik a cefrébe mártva méri annak hőmérsékletét, míg a másik légköri hőmérséklet méri.

Mamolar-Domenech és tsai. (2023) hőmérővel mérték a cefre hőmérsékletét ami a cefrébe merített hidrofon által mért hangokkal erősen korrelált de kisebb csúszással, ami a cefre hőelnyelő képességével magyarázható.

Az erjesztő helyiségben, hőmérővel folyamatosan mért hőmérséklet alapján kontrollálták Zhang és Wang (2013) a rizsbor cefréjének hőmérsékletét, egy izzólámpa ki-be kapcsolásával.

A erjesztőtálcákon (szilárd fázisú fermentáció), különböző pontokon mérték hőmérővel a koji hőmérsékletét Uehara és tsai. (2019), majd pedig, a folyamat további lépéseiben, a sake cefréjében, az erjesztőtartály alján, közepén és tetején (Uehara és tsai., 2018).

³⁹Kisebb cefre sűrűség = kisebb buborék méretek = magasabb frekvenciájú hangok.

⁴⁰<https://www.arduino.cc>

⁴¹Ezért, az ún. „zöld számítás” (https://en.wikipedia.org/wiki/Green_computing) alapelveit követve, egyre nagyobb jelentősége van a kis áram- és kis számításgényű technológiák, módszerek alkalmazásának.

⁴²<https://www.raspberrypi.com>

Vošahlík és Hart (2021) egy Raspberry PI⁴² és egy kompatibilis, folyadékszigetelt hőmérő segítségével mérték közvetlenül a rizsbor cefréjében annak a hőmérsékletét.

Andrade és tsai. (2021) szintén Raspberry PI⁴² segítségével gyűjtik a hőmérővel mért adatokat és vezérlik a házi sörfőzési folyamatot. Megjegyzik viszont, hogy ennek a technológiának a kompatibilitása, flexibilitása és megbízhatósága még nem éri el a nagyipari elvárásokat (mint pl. a PLC vezérlés technológiája), viszont kedvező ára miatt a kisipari és házi főzdek számára jó megoldás lehet.

Sainz és tsai. (2013) szabad hardveren (Texas Instruments⁴³) és szabad forráskódon alapuló rendszere az erjesztőtartályban mért hőmérsékletet vezeték nélküli technológiával továbbítja az adatfeldolgozó egységhez, pl. a borász telefonjára. Megjegyzik, hogy a vezeték nélküli technológiák nagy előnye az, hogy egyszerűbben telepíthetők és áthelyezhetők. A bemutatott szenzorral mért értékek megegyeztek a kísérletekben használt ipari hőmérővel⁴⁴) mért értékekkel.

Ranasinghe és tsai. (2013) a hőmérőknek az erjesztő tartályban való elhelyezésének fontosságát hangsúlyozzák. Egyrészt nem javasolják a fixen beépített szenzorokat, mivel ezek (a kiálló részek) könnyen megsérülhetnek tisztítás közben. Továbbá, megjegyzik, hogy nagy tartályok esetében több ponton is monitorozni kell a hőmérsékletkülönbségek miatt (főleg, ha a tartály nincs ellátva keverővel). Végül, nagy tartályok esetében, nem ritkán 5 méteres mélységben is monitorozni kell, így a szenzoroknak az itt fellépő nyomást is bírniuk kell. Ezért, egy 2.5 m hosszú cső alakú szondát fejlesztettek ki, amelyben 7 db hőmérő van egymástól 30-50 cm-es távolságra elhelyezve. Ezt a szondát, a tartály tisztításánál ki lehet emelni.

2.4.5. Savtartalom és pH

Tomtsis és tsai. (2016) „okos hordó” rendszerében a cefrébe merített digitális pH mérőt használnak.

Albanese és tsai. (2011) rendszerében laktát-oxidáz alapú bioszenzort (e-Nyelv) fejlesztettek ki a must tejsavtartalmának pontos mérésére.

Vošahlík és Hart (2021) egy Raspberry PI⁴² és egy kompatibilis pH mérő segítségével mérték közvetlenül a rizsbor cefréjének pH értékét.

2.4.6. Alkoholtartalom

Tomtsis és tsai. (2016) „okos hordó” rendszerében az erjesztőtérben elhelyezett MQ típusú gázszenzort (2.1. táblázat) használnak.

Az alkoholtartalom gázszenzorokkal történő mérését már régebben is alkalmazták.

⁴³<https://www.ti.com>

⁴⁴<https://www.hannainstruments.co.uk>

Austin és tsai. (1996) a cefrét teflon és szilikon membránok segítségével pervaporálták⁴⁵, hogy gáz halmazállapotú legyen, majd egy elektromos gázszenzor segítségével mérték a sör alkoholtartalmát.

Albanese és tsai. (2011) rendszerében alkohol-oxidáz alapú bioszenzort (e-Nyelv) fejlesztettek ki a must cukortartalmának pontos mérésére.

Calderon-Santoyo és tsai. (2010) egy kisméretű (2 literes) laboratóriumi bioreaktorból vezették el, egy automatizált szivattyú segítségével az erjesztőtérben lévő gázokat egy gázkromatográfba, aminek segítségével mérték a szintetikus „cefre” alkoholtartalmát.

Lidén és tsai. (2000) az erjesztőtérből kivezetett gázokban, e-Orr által mért adatokon és a cefrében, folyadékkromatográf által mért címkéken betanított neurális hálózatok segítségével predikálták, nagyon pontosan a cefre alkoholtartalmát.

Külső optikai szenzort használnak Cavinato és tsai. (1990) az alkoholtartalom mérésére, amely megfigyelő ablakkal rendelkező erjesztőtartályoknál működik. A fényt optikai kábel segítségével vezetik a világtótestből a tartályba és a tartályból az infravörös spektroszkópba, kikerülve a mintavételt ami a cefre szennyeződéséhez vezethet.

2.4.7. Aromaanyagok

Pinheiro és tsai. (2002) erjedő mustban, 20 napon át mérték az aromaanyagok fejlődését egy 32 gázszenzort tartalmazó kereskedelmi e-Orr-al. A cikkben bemutatott kísérletben kimutatták, hogy a csak etanolt tartalmazó minták és az ugyanolyan mennyiségű etanolt de még különböző hozzáadott aromaanyagokat is tartalmazó minták közti méréseknél lényegében ugyanazokat az „ujjlenyomatokat” kapták (a méréseknél a minták páratartalmát is egyformára állították be). Ezt azzal magyarázzák, hogy az aromakomponensek nagyon kis mennyiségben vannak jelen, amit a nagyobb mennyiségű etanol elfed a szenzorok elől. Ezért vagy az etanolt kell kivonni a mintákból vagy pedig az aromaanyagokat kell töményíteni. Mivel az alkohol kivonásával némely aromaanyagokat is kivonnánk a mustból, inkább annak pervaporációját⁴⁵ javasolják, pl. szilikon membrán segítségével. A kísérletekben kimutatták, hogy míg a pervaporált mintákban az etanol töménysége maximum tízszeresére nőtt, addig az aromakomponensek többségének a töménysége százszorosára nőtt. Következésképpen, a minták ujjlenyomatát már el lehetett választani egymástól és egyes szenzorok képesek voltak kimutatni az aromaanyagok fejlődésének menetét.

Szintetikus közegbe oltott, két fajta élesztő aromaképző képességét hasonlították össze Calderon-Santoyo és tsai. (2010). Az erjesztőtérből kivezetett és egy gázkromatográfba bevezetett gázokból, a gázkromatográf segítségével elválasztották a vizet és az alkoholt, és csak a gázokban lévő többi összetevőt vezették be az e-Orr-ba, egy szabadalmaztatott, ún. „back-flush” módszerrel.

⁴⁵A pervaporációnál a folyadék egyes alkotóelemei membránon áthaladva gáz halmazállapotba kerülnek.

Lidén és tsai. (2000) az erjesztőtérből kivezetett gázokban, e-Orr által mért adatokon és a cefrében, folyadékkromatográf által mért címkéken betanított neurális hálózatok segítségével predikálták a cefrében mért acetaldehid és etil-acetát tartalmát. Itt megjegyzik, hogy ezen alkotóelemek mennyisége nem érte el az e-Orr érzékelési küszöbét, ezért a neurális hálózatok (amelyek képesek komplex, nem lineáris összefüggéseket találni) jó útnak bizonyul ezen alkotóelemek közvetett módon való mérésére, amihez több adatra és kísérletre van szükség.

2.4. táblázat. Alkoholos erjesztés in-line és on-line monitorozását tárgyaló cikkek, jelezve a bennük bemutatott kutatás reprodukálhatóságát („R” oszlop) és a bemutatott módszerek megvalósíthatóságát („M” oszlop). Jelölések: „eO” = e-Orr, „eN” = e-Nyelv, „oSz” = optikai szenzor, „GC” = gázkromatográf, „eSz” = egyéb szenzor.

Publikáció	Monitorozás				R ⁴⁶	M ⁴⁷
	Közeg	Tulajdonság	Szenzor	Mód		
Albanese és tsai. (2011)	bor	% alk., fenolt., tejsav, glükóz	bio-eN	in-line	3	3
Andrade és tsai. (2021)	sör	°C	hőmérő	in-line	3	1
Austin és tsai. (1996)	sör	% alk.	eO	on-line	3	1
Becker és tsai. (2001)	sör	brix	ultrahang	on-line	3	2
Buonocore és tsai. (2021)	sör	CO ₂	mérleg	on-line	3	1
Calderon-Santoyo és tsai. (2010)	–	% alk., aroma	GC, eO	on-line	3	3
Cañete-Carmona és tsai. (2020)	bor	CO ₂	oSz	on-line	3	1
Cavinato és tsai. (1990)	–	% alk.	oSz	in-line	3	3
Lidén és tsai. (2000)	–	% alk., glükóz, CO ₂ , etilacetát, acetaldehid	eO, oSz	on-line	3	3
Lidgren és tsai. (2006)	–	glükóz	eSz	on-line	2	3
Mamolar-Domenech és tsai. (2023)	–	CO ₂	hidrofon	in-line	3	1
Pinheiro és tsai. (2002)	bor	aroma	eO	on-line	3	1
Ranasinghe és tsai. (2013)	bor	°C	hőmérő	in-line	3	2
Sainz és tsai. (2013)	bor	°C	hőmérő	in-line	3	1
Shrake és tsai. (2014)	bor	szín, fenolt.	oSz	in-line	3	3
Tamo és Hilario-Tacuri (2020)	sör	°C	hőmérő	on-line	3	1
Tomtsis és tsai. (2016)	bor	°C, pH, szín, CO ₂ , % alk.	hőmérő, pH, eO, szín, oSZ	on-line	2	2
Uehara és tsai. (2018)	sake	°C	hőmérő	in-line	3	1
Uehara és tsai. (2019)	koji	°C	hőmérő	in-line	3	1
Vasilescu és tsai. (2019)	bor	fenolt., glükóz	bio-eO	on-line	3	3
Vošahlík és Hart (2021)	rizsbor	°C, pH	pH, hőmérő	in-line	3	1
Zhang és Wang (2013)	rizsbor	°C	hőmérő	on-line	3	1

⁴⁶Három szintet különböztetünk meg: 1 = módszer és kísérletek hiányzó leírása, 2 = módszer vagy kísérletek hiányzó (hiányos) leírása, 3 = reprodukálható kutatás

⁴⁷Három szintet különböztetünk meg: 1 = otthon elkészíthető, 2 = jól felszerelt műhelyben elkészíthető, 3 = laboratóriumi körülmények közt készíthető el (pl. nanotechnológiák használata)

3. fejezet

Alkalmazott módszerek

Ahogy már említettem, dolgozatom áttekintő jellegű. A szakirodalmat a következő faktorok szerint elemeztem és összegeztem:

- A szenzorok típusa: A legfontosabb faktor, ami szerint a kutatásokat osztályozom, a szenzor típusa. Ez nem csak a gyűjtött adatok típusát definiálja, ami általában idősor, hanem azt is, hogy az adott méréseket hol végezzük el – a cefrében (pl. hőmérő, elektronikus nyelv) vagy a cefre feletti térben (pl. CO₂ szenzor, elektronikus orr), esetleg a tartályon kívül (pl. mérleg).
- Az erjesztendő alapanyag: A másik vizsgált faktor az erjesztendő alapanyag fajtája, ami meghatározza a cefre állagát is. Ez főként azoknál a szenzoroknál fontos kérdés, amelyek a cefrében mérik egy adott paraméter értékeit és esetleg a mérési pontosságukat befolyásolhatják a cefre paraméterei, mint pl. a pH vagy az alkoholtartalom. Ide kapcsolódik a cefre homogenitásának kérdése is. Egy keverővel ellátott erjesztőtartály esetében biztosak lehetünk abban, hogy az adott paraméter értéke a teljes cefrére vonatkozik, nem csak annak egy részére (pl. a cefre maghőmérséklete más mint a tartály falánál).
- Reprodukálhatóság és megvalósíthatóság Kutatóként fontosnak tartom, hogy a tudományos publikációk reprodukálhatók legyenek. Ehhez fontos, hogy az adott cikk érthetően és világosan tárgyalja a bemutatott kutatást, beleértve a felhasznált anyagok és módszerek pontos megnevezését, hivatkozásokkal együtt. Továbbá, fontos, hogy a cikkben bemutatott kísérletek és azok eredmények kiértékelése objektív legyen. Sajnos, a tudományos publikációk jelentős részénél ez nem így van (erre jutott Horváth és Carvalho, [2017](#) is több mint 80 cikket elemezve az informatika területéről), ezért ezeket a faktorokat is érinteni fogjuk. Végül pedig arra is kitérek, hogy a cikkekben bemutatott megoldások fizikai megvalósítása milyen szintű tudást és eszközöket igényel.

A kereséshez a Google[®] internetes böngészőt¹ használva. A Google[®] keresőjét azért választottam mert jó hatékonyságú és – tapasztalataim alapján – a különböző kiadók (pl. Elsevier, Springer, MDPI, Wiley, stb.) adatbázisait is megfelelően indexeli.

A szakirodalom feldolgozását, „felülről lefelé haladva” (úm. az általánosabb cikkektől kiindulva, a kimondottan alkoholos erjesztéssel foglalkozó cikkekig), a következő lépésekben végeztem:

- Survey cikkek: Először az ún. „survey” cikkeket gyűjtöttem össze, amelyek a szenzorok élelmiszeriparban való alkalmazhatóságának szakirodalmát térképezik fel (2.3. fejezet). Itt a szenzorok élelmiszeripari – tehát nem csak szeszipari – felhasználására irányuló cikkekre fókuszáltam a nagyobb lefedettség érdekében. A jó survey cikkekben, amelyeket az adott terület szakértői írtak, a szakirodalom nem csak egyszerűen listázva van hanem rendszerezve is, különböző elvek alapján, feltárva az adott szakirodalomban lévő hiányosságokat. Ezért találtam hasznosnak a survey cikkeket a kutatás kiindulási alapjaként.
- Kiindulási publikációk: A második lépésben azokat a publikációkat szűrtem ki a survey cikkekben tárgyaltak közül, amelyek kimondottan a szenzorok szeszipari felhasználásához kapcsolódnak. Ehhez a listához (következőkben $\mathcal{P}$ -vel jelölve) még hozzáadtam azokat a publikációkat amelyeket a Google[®] böngésző megtalált de nem voltak a survey cikkekben megemlítve (2.4. fejezet).
- Elemzés: A harmadik fázisban a következő, ún. szélességi bejárás² szerint elemeztem a cikkeket illetve vontam be az elemzéshe újabbakat (2.4. fejezet):
 1. A $\mathcal{P}$ listából elemezd a – sorrendben első – p publikációt a fent említett faktorok alapján (szenzor típusa, alapanyag, adatelemzés módszerei, reprodukálhatóság és objektivitás, illetve megvalósíthatóság);
 2. A p -ben idézett, releváns publikációkat $(p_1, p_2, \dots, p_n)$ csatold a $\mathcal{P}$ végére;
 3. Töröld p -t a $\mathcal{P}$ elejéről;
 4. Ha $\mathcal{P}$ nem üres akkor ugorj az 1. pontra, ellenkező esetben vége;

¹<https://www.google.com>

²https://hu.wikipedia.org/wiki/Szélességi_bejárás

4. fejezet

Eredmények és értékelésük

A szakirodalom áttekintése alapján (2. fejezet) állíthatom, hogy a gyümölcscefrék erjesztésének automatizált monitorozásával, sajnos, nem foglalkozik a szakma, ellentétben a lepárlóberendezések automatizálására tett törekvésekkel. Ez valószínűleg azért van, mert a legtöbb és világon a legnagyobb mértékben gyártott szeszesital (whisky, vodka, rum, cachaça, tequila, cognac) folyékony cefréjének erjesztéséhez könnyen alkalmazhatók a borászatban és a sörgyártásnál alkalmazott erjesztést monitorozó technológiák.

A másik ok az lehet, hogy ezekhez az italokhoz képest, „kicsi a szakma”, vagyis a gyümölcspárlatok készítésének, illetve azok fogyasztásának a hagyománya egy-két orszégra terjed csak ki, mint pl. a német nyelvterületeken ismert schnapps vagy obstwasser, és a Kárpát-medencei pálinka (Magyarország), pálenka (Szlovákia és Csehország) vagy palincă (Románia). A minőségi gyümölcspárlat- és pálinkakészítés többnyire családi- vagy kisvállalkozásokban történik és szezonális jellegű, így az erjesztés monitorozását többnyire off-line módon, emberi munkaerő végzi.

4.1. Értékelés szakmai szempontok alapján

A szakirodalomban található megoldásokat a főzde és főzőmester szempontjából elemzem ebben a részben.

Megtérülés Egy pálinkafőzde szempontjából, üzleti vállalkozás révén, fontos kérdés az erjesztés automatikus monitorozásába fektetett tőke megtérülése. Ausztrál borászatokban végzett kutatás alapján, Scrimgeour (2000) jelentése szerint, a must szárazanyagtartalmának mérésére való ipari szenzor technológiák kis borászatokban¹ való telepítésének ára 22.500\$AUS, aminek a megtérülése kb. 3 év. Vasilescu és tsai. (2019) monitorozó rendszerének az ára is hasonló nagyságrendű. (A megtérülést illetően, még érdemes megjegyezni, hogy az utóbbi évek időjárási viszonyait tekintve a borászatok nagyobb biztonsággal² is

¹Átlagosan, 10 db 7.500 literes erjesztőtartállyal, évi 2.250 hektoliter mustot erjesztve.

²Pl. az utóbbi öt évben nem termelt elegendő sárgabarack Magyarországon (a fagykárak miatt).

tervezhetnek mint a pálinkafőzdék.) Ehhez kell még számolni a fenntartás (szenzorok tisztítása, kalibrálása, stb.) költségeit, amik már lényegesen kisebbek a telepítési költségeknél. A fent említett megtérülési idő a kis (főként a Kárpát-medencei) pálinkafőzdéknél ennek a többszöröse is lehet. Egyenlőre még nem találni a piacon elfogadható áron, és kimondottan cefre erjesztéséhez készített bioszenzorokat.

Megbízhatóság Az előző bekezdés alapján, az olcsó és több erjesztési paramétert egyszerre monitorozó módszerek jelenthetnek megoldást a pálinkafőzdék számára. Ezek, viszont, egyenlőre még inkább csak laboratóriumi környezetben lettek tesztelve. A bioszenzorok esetében kevés adat is áll rendelkezésre azok stabilitásáról és élettartamáról, ami lehet nagyon rövid (kb. 1 hét az etanol szenzornál) vagy hosszú (több hónap a glükóz szenzornál) – ezeket az időtartamokat mindenképpen ipari környezetben kell tesztelni ahhoz, hogy biztonságosan meg tudjuk állapítani azt, hogy melyik szenzor, milyen cefrénél, mikor szorul cserére. Az áttekintésből (2.4. fejezet) látszik, hogy sok kutatásban Raspberry PI-t vagy Arduino-t használtak a monitorozás vezérlésére. Ezek a technológiák nem ipari környezethez, hanem inkább edukációs célokra lettek kifejlesztve, ezért nem biztos, hogy az erjesztőhelyiség környezete (pl. a páratartalom, nedvesség, tisztítószer, stb.) nem hat ki negatívan ezek hosszú távú működésére. Végül pedig, kimondottan gyümölcscefre erjesztését monitorozó megoldásokat nem találtam. Ilyen jellegű kutatások és tesztek szükségesek ahhoz, hogy a sörgyártásból és a borászatból átvett módszerek megbízhatóan alkalmazhatók legyenek a gyümölcscefre erjesztésének monitorozását illetően. Scrimgeour (2000) megjegyzi, hogy az általuk tesztelt, borászatokba telepített ipari szárazanyag-tartalmat monitorozó rendszerek nem voltak teljesen megbízhatóak és pontosságuk elmaradt a referenciamódszerekétől³.

Ergonómia Az erjesztést monitorozó módszereknél ügyelni kell arra, hogy használatuk biztonságos legyen fizikai és kémiai szempontból egyaránt, a balesetek és a szennyeződések elkerülése végett. Az áttekintésben bemutatott módszerek nagy részét nem tesztelték ilyen szempontból ipari környezetben. A legtöbb bemutatott módszer telepítéséhez a tartályokat módosítani kell, pl. csap beépítésével az automatikus mintavevőhöz (2.1. ábra). Ezen módszerek egyes moduljai a tartályban, a tartályon, vagy a tartály mellett helyezkednek el, amik sérülhetnek a tartályok tisztítása közben. Egy másik kérdés, amit szintén nem teszteltek a bemutatott módszereknél ipari környezetben, hogy a szenzorok cseréje alatt mennyire van kitéve szennyeződésnek a cefre. Ez fontos kérdés, hiszen egyes bioszenzorok élettartama kisebb mint a gyümölcscefrék erjedési ideje. A pálinkafőzdéknél több típusú és méretű erjesztőtartályt használnak, a kis méretű műanyag hordóktól vagy úszófedeles acéltartályoktól (kis mennyiségű, pl. vadontermő gyümölcsök esetében), a

³Szesesitaloknál lásd pl. az Eurofins Food and Feed Testing Budapest Kft. hatályos, akkreditációs okiratában (<https://nah.gov.hu/hu/szervezet/eurofins-food-and-feed-testing-budapest-kft-elelmi/>).

pár ezer literes hűtőköpenyes tartályokon keresztül (ez az általános típus), a több ezer literes, hűtőköpennyel, keverővel és CIP rendszerrel ellátott tartályokig (inkább a nagyobb főzdeknél találhatók meg). Sok cikkből hiányzott az az információ, hogy milyen tartályok esetében használhatók, pl. hogy a tartályban lévő nyomás, a törkölykalap, vagy éppen a cefre keverése mennyire befolyásolja a szenzorok működését.

4.2. Értékelés kutatói szemszögből

A szakirodalomba található kutatásokat elemzem a további kutatások szemszögéből ebben a részben.

Reprodukálhatóság, Megvalósíthatóság és Objektivitás A 2.4. táblázat hatodik és hetedik oszlopából látszik, hogy a bemutatott módszerek jól reprodukálhatók és megvalósíthatók, bizonyos szakértői tudást és eszköztárat igényelve. Az elemzett cikkek objektivitását tekintve vannak kifogásaim: Jól látszik, és a cikkek szerzői is sokszor megemlítik, hogy hiányzik egy egységes és szabadon hozzáférhető adatbázis. Viszont, egyik cikkben sem találtam információt arról, hogy az adatokat közzé tették volna a szerzők, annak ellenére, hogy rengeteg olyan platform létezik, amin ezt könnyen meg lehet oldani, de egy weboldalra feltöltött fájl is elegendő lenne. A bemutatott módszerek tesztelésénél csak egy-két típusú bort, sört, esetleg szintetikus közeget használtak. Ennél többre lenne szükség ahhoz, hogy pontosabb képet kapjunk a bemutatott módszerek és szenzorok tulajdonságairól (nem hinném, hogy gond lenne sokféle tulajdonságú és különböző szeszesitalok cefréjéhez hasonló szintetikus közeget összeállítani). Végül pedig, azt látom, hogy sok kis elszigetelt kutatásról van szó, amelyek nem hasonlítják össze a bemutatott módszert más módszerekkel, annak ellenére sem, hogy azok egyébként jól dokumentáltak (pl. a bioszenzorokat bemutató cikkekben nagyon pontosan le van írva azok elkészítésének menete). Így lehetetlen összehasonlítani két különböző, de ugyanazt a tulajdonságot monitorozó módszert.

Gyümölcscefrék erjedésének monitorozása A gyümölcsök cefréi nem annyira egységesek mint a többi szeszesital cefréi. A gyümölcscefre lehet folyékonyabb (pl. málna), pépesebb (pl. szilva) és nagyon sűrű (pl. birs, törköly), aminek a halmazállapota, inhomogenitása és viszkozitása negatívan hathat a borok és sörök cefréjénél használt szűrő membránok működésére (pl. az e-Nyelv esetében). Továbbá, a gyümölcsök aromaanyagai és az erjedésnél keletkező anyagok (mint pl. a meggy esetében keletkező magasabb kén-tartalom) heterogenitása megnehezítheti a monitorozást, vagy éppenséggel el is fedheti a nagyon kis mértékben jelenlévő anyagokat. Ahhoz, hogy a bemutatott megoldásokat a gyümölcscefék monitorozásánál is használhassuk, el kellene kezdeni a kísérleteket különböző szenzorokkal és különböző gyümölcsfajtákkal (kezdve a legáltalánosabbakkal, pl. alma,

körte, szilva). Ilyen kísérletek és azok eredményeinek elemzése nélkül bármilyen, más típusú szeszesital cefréjénél használt erjesztést monitorozó megoldások nem használhatók biztonságosan. Az egyedüli, pépesebb cefrét monitorozó megoldások a rizsből készített szeszesitalokhoz kapcsolódnak és csak a cefre hőmérsékletét mérik, egy kivétellel (Vošahlík és Hart, 2021) ahol a cefre pH-ját mérik sztenderd pH szenzorral. Gyümölcscefrék erjedését monitorozó megoldást nem találtam a szakirodalomban.

4.2.1. Adatok elemzése

Az áttekintett cikkekben az adatok elemzésére használt módszerek az adattudomány és a GT széles palettájának (2.2. táblázat) két végletébe tartoznak: vagy a nagyon egyszerű, statisztikai módszereket, vagy pedig a nehezen interpretálható módszereket alkalmazzák, úm. faktor analízist vagy neurális hálózatokat, amelyek egyre népszerűbbek az élelmiszeripari alkalmazásoknál (Nyitrai és tsai., 2018). Több cikk is alkalmaz neurális hálózatokat a szenzorok által generált adatok elemzésére, de sajnos a legtöbb esetben nem használták, vagy pedig nem tesztelték szakszerűen ezeket a modelleket, amit a következő bekezdésekben fejtek ki, két-három cikken illusztrálva a legfőbb problémákat és kihívásokat.

Komplex modellek betanítása kis adathalmazokon A legtöbb cikkben nagyon kis mennyiségű adatot gyűjtöttek a bemutatott szenzorokkal. Ez még nem lenne baj, viszont a GT modell kiválasztásánál az egyik fontos faktor az, hogy az adathalmaznak megfelelő modelleket használjunk, nem kell mindenáron a neurális hálózatokat erőltetni. Qi és tsai. (2017b) hét baiju-ból vett 30-30 mintából álló adathalmazt, azaz 210 idősortanítottak be egy mély konvolúciós neuronhálót. Hasonlóan komplex modellt használtak Zhang és tsai. (2021) 10 baiju-ból vett 6-6 mintán, azaz 60 idősortan. Ezekkel ellentétben, Lidén és tsai. (2000) egy egyszerű, egy rejtett réteggel rendelkező neurális hálózatot használtak hét mérésből származó, 10 szenzor által gyűjtött idősorokból kinyert 105 jellemzőre (tehát nem közvetlenül az idősorokra), ami egy korrekt választásnak tűnik. Itt még meg kell jegyeznem, hogy kis adathalmazokat lehet „augmentálni”, azaz a létező adatokból sok új szintetikus adatot generálni⁴, viszont a szenzorok által generált adatok esetében, méginkább ha nagyon kicsi eltérések vannak az idősorokban, ez problémákat is okozhat.

Hiányzó vagy pedig nem elegendő hiper-paraméter hangolás Egy GT algoritmusnak, illetve az optimalizált modellnek vannak paraméterei (hiper-paraméterek) amiket a tanítás előtt be kell állítani. Ezt különböző módszerekkel végezhetjük, amik a legtöbb GT szoftverben megtalálhatók. Persze, használhatjuk az algoritmusok alap beállításait is, amik sokszor jól működnek (Horváth és tsai., 2023), viszont ajánlatos finomhangol-

⁴Pl. egy képet el tudunk forgatni 359 különböző fokkal, meg tudjuk picit nyújtani hosszában vagy széltében, stb. több száz új, hasonló, de mégis picit más képhez jutva.

ni ezeket⁵. Ennek ellenére, Qi és tsai. (2017b) egyáltalán nem végeztek finomhangolást, míg Zhang és tsai. (2021) csak a tanító algoritmus hiper-paramétereit hangolták, a neurális hálózat hiper-paramétereit nem. Lidén és tsai. (2000) végeztek valamilyen szintű hiper-paraméter hangolást.

Hiányzó „baseline” modellek Ahhoz, hogy eldöntsük, hogy melyik GT modellt válasszuk az adatainkra, érdemes több modellt is összehasonlítani. Ennek megvannak az elfogadott módszerei (ezek, lényegében a statisztikában jól ismert hipotézis tesztelő módszerek). Lidén és tsai. (2000) nem használtak baseline módszereket, Qi és tsai. (2017b) és Zhang és tsai. (2021) használtak baseline modelleket de azok hiper-paramétereit már nem hangolták be, ami nem becsületes összehasonlítást eredményez.

„Feature engineering” A nyers adatok sok olyan jellemzőt (feature) vagy információt tartalmazhatnak, amiket felhasználva, még a viszonylag egyszerű GT modellel is jó teljesítményt érhetünk el. Ezeknek a kinyerését nevezzük feature engineeringnek⁶, amit manapság már, sajnos, nem nagyon alkalmazunk. Qi és tsai. (2017b) és Zhang és tsai. (2021) nem alkalmaztak feature engineering-et, míg Lidén és tsai. (2000) igen, aminek eredményeként egy egyszerű modellel is nagyon jó predikciós teljesítményt értek el.

„Zöld számítás” A világ energiafogyasztásának kb. 5-10%-áért felelősek a számítógépek, aminek jelentős részét a mesterséges intelligencia (köztük a GT) modelljeinek a betanítása és használata képez. Ezért indult el a 2. fejezetben már említett zöld számítás⁷ néven futó kezdeményezés, amely a fenntartható számításokat szorgalmazza. Az én értelmezésemben ez azt jelenti, hogy ha lehet, akkor olyan GT modelleket használjunk, amiknek betanítása és használata kis számításigényű. Ezek főként az egyszerű, pár száz vagy pár ezer paraméterből álló modellek, tehát kompromisszumokat kell találnunk a modellek komplexitása és pontossága között. Ezt azzal érhetjük el, hogy feature engineeringet használunk és az egyszerű modelleknél is alapos hiper-paraméter hangolást végzünk (Lidén és tsai., 2000). Az egyszerű modellek egyszerű hardveren is telepíthetők és futtathatók, ami az adott módszer árában is tükröződni fog.

Interpretálható modellek A erjedés monitorozása közben, szenzorok által gyűjtött adatok elemzésének (a pontos predikciók vagy anomália detektálás mellett) az is a célja kellene, hogy legyen, hogy az adatokban lévő, rejtett információkat és mintázatokat

⁵Hasonlóan, ahogyan a pálinkafőző berendezés paramétereit (pl. a tányérok magassága, deflegmáció és a fűtés intenzitása) is az adott cefrének és a főzés menetének megfelelően állítjuk be.

⁶A FA (köztük a PCA) is egyfajta feature engineering módszer amit az élelmiszeriparban előszeretettel alkalmaznak.

⁷https://en.wikipedia.org/wiki/Green_computing

a felszínre hozzuk. Ezt azzal érhetjük el a legjobban, ha olyan modelleket is alkalmazunk, amik az ember által könnyen értelmezhetők (2.2. táblázat). Sajnos, az áttekintett szakirodalomban erre nem találtam példát.

Benchmark adathalmazok, meta-adatok és adatvédelem Amint azt több cikkben is megemlítik (2.3.2. fejezet), nagy szükség lenne egy szabadon elérhető adatbázisra, amely nemcsak a szenzorok által gyűjtött nyers adatokat tartalmazza idősorok formájában, hanem a monitorozás adatait is (mikor, milyen cefrében, milyen szenzorokkal lettek az adatok gyűjtve) – ezeket nevezzük meta-adatoknak. Itt viszont tudatosítanunk kell, hogy egy pálinkafőzdében monitorozott erjesztés szenzorok által gyűjtött adatai, valamint azok meta-adatai érzékeny adatok és üzleti titoknak minősül(het)nek. Így fontos az adatok titkosításának és anonimizálásának a biztosítása. Ennek hiányában nem lehetséges nagy mennyiségű, megbízható adatot gyűjteni.

5. fejezet

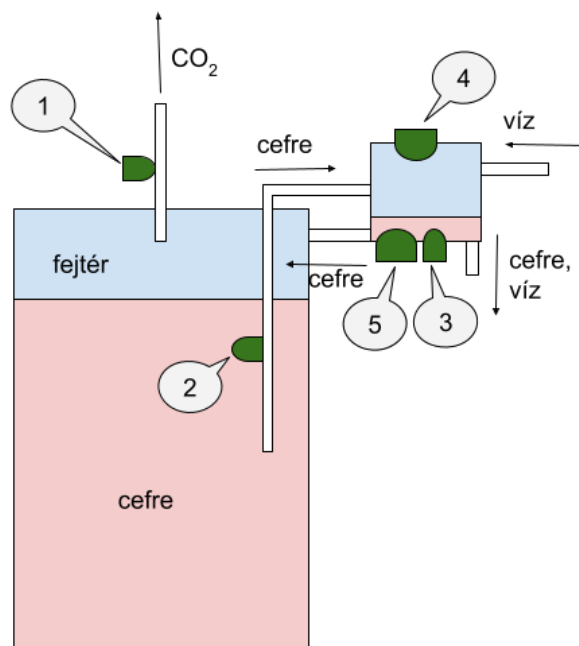
Következtetések és javaslatok

Az előző két fejezetben felvázolt ismeretek alapján, a gyümölcscefrék erjesztésének automatizált monitorozásának kérdésében, az alábbi következtetéseket tudom levonni:

- A cefre legfontosabb fizikai (hőmérséklet) és kémiai (pH) paramétereit az eddig is használt, elektronikus (alkoholtűrő) folyadékhőmérővel és elektronikus pH mérővel mérhetjük, a szenzorokat közvetlenül a cefrében elhelyezve (Vošahlík és Hart, 2021).
- A cefre további paramétereit közül a CO₂ képződést érdemes mérni, legegyszerűbben a kotyogóra szerelt fotodiódával, egységnyi idő alatt áthaladó buborékok számát számolva (Tomtsis és tsai., 2016). Mivel a CO₂ képződés arányos az erjeszthető cukortartalom csökkenésével és az alkoholtartalom növekedésével, így ez utóbbi két paraméter levezethető az előbbiből.
- Az elektrokémiai e-Orr és e-Nyelv használatát illetően nincsenek a szakirodalomban nagyobb adatokon végzett, objektív kísérletekkel alátámasztott konkluzív eredmények. Abban azonban megegyezés van, hogy az alkohol- és CO₂ tartalom nincs jó hatással a szenzorokra.
- A bioszenzorok megoldást jelenthetnek a cefre több paraméterének a mérésére, azonban ezek a módszerek még inkább csak laboratóriumi környezetben lettek tesztelve, alkalmazásaik nem terjedtek még el a szesziparban.
- Az erjesztést folyamatosan monitorozó megoldások nem voltak még eddig gyümölcscefréken tesztelve, így azok alkalmazhatóságát legalább laboratóriumi környezetben tesztelni kell.
- A szoft (főként a GT alapú) szenzorok fejlesztéséhez nem áll rendelkezésre elegendő mennyiségű adat. Alapos adatgyűjtéssel és adatfeldolgozással, interpretálható modellek segítségével, azonosíthatóak lehetnének azok a paraméterek amiket szükséges mérni, amik alapján egyszerűbb és olcsóbb megoldásokat lehetne fejleszteni (pl. a redundáns szenzorok használatának kiszűrésével).

A fenti következtetések alapján a következő, moduláris felépítésű erjedést monitorozó rendszer vázlatát javaslom, amelyet a felhasználó a saját igényei szerint állíthat össze.

5.1. ábra. A javasolt erjedést monitorozó rendszer leegyszerűsített vázlata



Az erjesztőtartályon lévő kotyogóra van felszerelve a fotodiódás, buborékokat számláló CO₂ mérő (1) (Tomtsis és tsai., 2016). Ez az alap modulja a rendszernek, ami a CO₂ képződés dinamikája alapján jelezni tudja, ha elakad az erjedés és a főzőmester beavatkozása szükséges.

A második modul a hőmérő (2), ami attól függően, hogy van-e rá szükség¹, a cefrébe van helyezve, a szívócsőben, vagy arra erősítve (lásd. lejjebb) (Ranasinghe és tsai., 2013). A CO₂ képződés és a hőmérséklet azok a paraméterek amiket jó nagyon gyakran, vagy folyamatosan monitoroznunk.

A rendszer többi moduljához egy kis, zárt minta tartály szükséges, ami az erjesztőtartályra van erősítve. A mintavevés beállított időközönként történik. Ebbe vezetjük be egy kis szivattyúval a szűrt cefrét egy csövön keresztül (a cefrébe lógatott cső végén van a szűrő). A minta tartály aljába beszerelhető a pH mérő (3) és az e-Nyelv (5), a főzőmester döntése alapján. A minta tartály tetején beszerelhető az e-Orr (4), szintén igény szerint.

A mérések után a cefrét a minta tartályból elvezetjük, vagy vissza az erjesztő tartályba (kis mennyiségű cefre esetén²), vagy pedig a lefolyóba. A minta tartályt ezután tiszta vízzel kimossuk, amit elvezetünk a lefolyóba.

A rendszert, aminek a része a szivattyú és a minta tartályba be- és abból kivezető csöveken lévő szelepek, egy Raspberry PI-vel vezérelhetjük,

¹Hűthető/fűthető tartály esetében nem szükséges hőmérőt beszerelni.

²Mivel a minta tartály zárt, így a szennyeződés lehetősége nem áll fenn

6. fejezet

Összefoglalás

Hosszútávú tervem, hogy egy olyan erjedést monitorozó rendszert fejlesszek, amely a főzőmester igényei szerint állítható össze. Adattudománnyal és gépi tanulással foglalkozom, így természetesen a célom, hogy ehhez a rendszerhez egy adatelemző modul is tartozzon.

Ennek a hosszútávú tervnek, természetszerűen, az első lépése, hogy feltérképezem a létező megoldásokat, kielemezzem azt, hogy milyen típusú szenzorokkal tudjuk mérni a cefre és az erjedés egyes paramétereit és betekintést nyerjek abba, hogy milyen módszerekkel elemzik az adatokat az adott szakterületen.

Dolgozatomban az erjesztés monitorozásáról szóló kutatásokat tekintetem át, a bio- és gyógyszeripari alkalmazásoktól, az élelmiszeripari alkalmazásokon keresztül, a szeszipari alkalmazásokig. Sajnos, nagyon kevés cikket találtam ami kimondottan a szeszesitalok erjesztésének on-line vagy in-line monitorozásával foglalkozik, és egyet sem ami a gyümölcscefre erjesztéssel kapcsolatos.

A sok lehetséges irány közül terjedelmes szakirodalma van az e-Orr és az e-Nyelv alkalmazásainak, illetve az ezekkel a technológiákkal kapcsolatos bioszenzorok fejlesztésének. Sajnos, a nagy mennyiségű, és jól reprodukálható szakcikk ellenére azt látom, hogy hiányoznak a közös és elérhető adatbázisok és az egységes sztenderdek, amik alapján konkluzív eredményekre egyelőre nem számíthatunk. Ez viszont, annak ellenére, hogy hátráltatja az erjesztés monitorozására irányuló ipari fejlesztéseket, lehetőségeket is hordoz magában. Egyrészt egy gyümölcscefrék erjesztését monitorozó rendszer fejlesztésére irányuló kutatás beindítását, nagyszabású adatgyűjtés megkezdését és az ezekre a típusú adatokra szabott mesterséges intelligencia modellek fejlesztését.

Ez nem egy rövidtávú kutatás és fejlesztés, több szakterület bevonásával (pálinkamesterek, élelmiszeripari, adatgazdálkodási és mesterséges intelligencia kutatók), viszont van értelme belekezdeni, még akkor is ha a pálinka és a gyümölcspárlatok „világa” nagyon kicsi, a világ vezető szeszesitalaihoz viszonyítva.

Remélem, hogy dolgozatom hasznos információkkal szolgál ehhez a munkához azoknak akik ebben az irányban szeretnének kutatni a jövőben.

Köszönetnyilvánítás

Először is, szeretném megköszönni témavezetőimnek, Dr. Kun Szilárdnak és Dr. Sipos Lászlónak a hasznos tanácsokat a dolgozat elkészítéséhez. Remélem, hogy a jövőben is együtt tudunk majd dolgozni.

Nagyon köszönöm feleségem, Danielle Alessandra támogatását a tanulmányaim alatt és a dolgozat elkészítése közben. Kislányom, Sofia Lilla sokat nélkülözött engem ez alatt a négy szemeszter alatt, remélem kárpótolni tudom majd a nem együtt töltött időért.

A pálinkához fűződő legkedvesebb emlékeim azok a napok, amikor apámmal pálinkát főztünk. Ültünk, reggeltől estig az üst mellett, raktuk a tüzet, beszélgettünk, és néztük, hogyan csordogál a pálinka. Ezeknek a napoknak volt egyfajta ünnepi hangulata. Apa már nincs köztünk de nélküle nem biztos, hogy megszületett volna bennem az érdeklődés és a kedv a pálinkakészítés mesterségének a tanulásához. Az Ő emlékének ajánlom ezt a dolgozatot.

Irodalomjegyzék

- Abrankó, László (2023). *Pálinka-minőségvizsgálat*. Előadás anyagai (diasorok). Magyar Agrár- és Élettudományi Egyetem, Pálinkamester szaktanácsadó szakirányú továbbképzési szak.
- Adeleke, Ismail, Mnamdi Nwulu és Oluwafemi Ayodeji Adebo (2023). “Internet of Things (IoT) in the food fermentation process: A bibliometric review”. *Journal of Food Process Engineering* 46.5, e14321. DOI: [10.1111/jfpe.14321](https://doi.org/10.1111/jfpe.14321).
- Akin, Mehriban, Merve Yuksel, Caner Geyik, Dilek Odaci, A. Bluma, Tim Höpfner, Sascha Beutel, Thomas Scheper és Suna Timur (2010). “Alcohol biosensing by polyamidoamine (PAMAM)/cysteamine/alcohol oxidase-modified gold electrode”. *Biotechnology Progress* 26. DOI: [10.1002/btpr.372](https://doi.org/10.1002/btpr.372).
- Albanese, Donatella, Consolatina Liguori, Vincenzo Paciello és Antonio Pietrosanto (2011). “Winemaking Process Monitoring Based on a Biosensor Automatic System”. *IEEE Transactions on Instrumentation and Measurement* 60.5, 1909–1916. old. DOI: [10.1109/TIM.2010.2102410](https://doi.org/10.1109/TIM.2010.2102410).
- Aleixandre, M., J. Lozano, J. Gutiérrez, I. Sayago, M.J. Fernández és M.C. Horrillo (2008). “Portable e-nose to classify different kinds of wine”. *Sensors and Actuators B: Chemical* 131.1, 71–76. old. DOI: [10.1016/j.snb.2007.12.027](https://doi.org/10.1016/j.snb.2007.12.027).
- Alexy, Márta és Tomáš Horváth (2022). “Tracing the Local Breeds in an Outdoor System – A Hungarian Example with Mangalica Pig Breed”. *Tracing the Domestic Pig*. Szerk. Dr. Goran Kušec és Dr. Ivona Djurkin Kušec. Rijeka, Croatia: IntechOpen. 5. fej. DOI: [10.5772/intechopen.101615](https://doi.org/10.5772/intechopen.101615).
- Allepuz Rico, Xavier (2020). *A low-cost multispectral camera*. BSc. Thesis. Universitat de Barcelona.
- Alpaydin, Ethem (2010). *Introduction to Machine Learning*. 2nd. Cambridge, Massachusetts, US: The MIT Press. DOI: [10.1017/S1351324912000290](https://doi.org/10.1017/S1351324912000290).
- Andrade, Alexandre Acácio de, Álvaro Batista Dietrich, Júlio Francisco Blumetti Facó és Ricardo Reolon Jorge (2021). “Low Cost Solution for Home Brewing and Small Brewing Business Using Raspberry Pi”. *Proceedings of IDEAS 2019*. Cham: Springer International Publishing, 136–145. old. DOI: [10.1007/978-3-030-55374-6_14](https://doi.org/10.1007/978-3-030-55374-6_14).

- Austin, Glen D., Inge Russell, Anton G. Metering és Ronald E. Subden (1996). “A Gas-Sensor-Based On-Line Ethanol Meter for Breweries”. *Journal of the American Society of Brewing Chemists* 54.4, 212–215. old. DOI: [10.1094/ASBCJ-54-0212](https://doi.org/10.1094/ASBCJ-54-0212).
- Baldwin, Elizabeth A., Jinhe Bai, Anne Plotto és Sharon Dea (2011). “Electronic Noses and Tongues: Applications for the Food and Pharmaceutical Industries”. *Sensors (Basel, Switzerland)* 11, 4744–4766. old. DOI: [10.3390/s110504744](https://doi.org/10.3390/s110504744).
- Becker, T., M. Mitzscherling és A. Delgado (2001). “Ultrasonic Velocity – A Noninvasive Method for the Determination of Density during Beer Fermentation”. *Engineering in Life Sciences* 1.2, 61–67. old. DOI: [10.1002/1618-2863\(200108\)1:2<61::AID-ELSC61>3.0.CO;2-D](https://doi.org/10.1002/1618-2863(200108)1:2<61::AID-ELSC61>3.0.CO;2-D).
- Bergmeir, Christoph és José M. Benítez (2012). “On the use of cross-validation for time series predictor evaluation”. *Information Sciences* 191, 192–213. old. DOI: [10.1016/j.ins.2011.12.028](https://doi.org/10.1016/j.ins.2011.12.028).
- Biechele, Philipp, Christoph Busse, Dörte Solle, Thomas Scheper és Kenneth Reardon (2015). “Sensor systems for bioprocess monitoring”. *Engineering in Life Sciences* 15.5, 469–488. old. DOI: [10.1002/elsc.201500014](https://doi.org/10.1002/elsc.201500014).
- Blanco, Carlos A., Rocío de la Fuente, Isabel Caballero és María L. Rodríguez-Méndez (2015). “Beer discrimination using a portable electronic tongue based on screen-printed electrodes”. *Journal of Food Engineering* 157, 57–62. old. DOI: [10.1016/j.jfoodeng.2015.02.018](https://doi.org/10.1016/j.jfoodeng.2015.02.018).
- Branchini, Carla Guanais, Larisa Lvova, Corrado Di Natale és Roberto Paolesse (2016). “Chapter 29 - Wine and Combined Electronic Nose and Tongue”. *Electronic Noses and Tongues in Food Science*. Szerk. María Luz Rodríguez Méndez. San Diego: Academic Press, 301–307. old. DOI: [10.1016/B978-0-12-800243-8.00029-9](https://doi.org/10.1016/B978-0-12-800243-8.00029-9).
- Buonocore, Daniele, Giuseppe Ciavolino, Domenico Di Caro és Consolatina Liguori (2021). “An IoT-based beer fermentation monitoring system”. *2021 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor)*. ACM, 263–267. old. DOI: [10.1109/MetroAgriFor52389.2021.9628536](https://doi.org/10.1109/MetroAgriFor52389.2021.9628536).
- Buratti, S., D. Ballabio, G. Giovanelli, C.M. Zuluanga Dominguez, A. Moles, S. Benedetti és N. Sinelli (2011). “Monitoring of alcoholic fermentation using near infrared and mid infrared spectroscopies combined with electronic nose and electronic tongue”. *Analytica Chimica Acta* 697.1, 67–74. old. DOI: [10.1016/j.aca.2011.04.020](https://doi.org/10.1016/j.aca.2011.04.020).
- Buratti, Susanna és Simona Benedetti (2016). “Chapter 28 - Alcoholic Fermentation Using Electronic Nose and Electronic Tongue”. *Electronic Noses and Tongues in Food Science*. Szerk. María Luz Rodríguez Méndez. San Diego: Academic Press, 291–299. old. DOI: [10.1016/B978-0-12-800243-8.00028-7](https://doi.org/10.1016/B978-0-12-800243-8.00028-7).
- Cabanes, Francisco Javier, Natasha Sahgal, M. Rosa Bragulat és Naresh Magan (2009). “Early discrimination of fungal species responsible of ochratoxin A contamination of

- wine and other grape products using an electronic nose”. *Mycotoxin Research* 25.4, 187–192. old. DOI: [10.1007/s12550-009-0027-x](https://doi.org/10.1007/s12550-009-0027-x).
- Calderon-Santoyo, Montserrat, Pascale Chali r, Dominique Chevalier-Lucia, Charles Ghommidh  s Juan Arturo Ragazzo-Sanchez (2010). “Identification of *Saccharomyces cerevisiae* strains for alcoholic fermentation by discriminant factorial analysis on electronic nose signals”. *Electronic Journal of Biotechnology* 13, 8 –9. old. DOI: [10.2225/vol13-issue4-fulltext-11](https://doi.org/10.2225/vol13-issue4-fulltext-11).
- Calvo-L pez, Antonio, Oriol Ymbern, David Izquierdo  s Juli n Alonso-Chamarro (2016). “Low cost and compact analytical microsystem for carbon dioxide determination in production processes of wine and beer”. *Analytica Chimica Acta* 931, 64–69. old. DOI: [10.1016/j.aca.2016.05.010](https://doi.org/10.1016/j.aca.2016.05.010).
- Cargalo, Carina L., Pau Cabaneros Lopez, Aliyeh Hasanzadeh, Isuru A. Udugama  s Krist V. Gernaey (2022). “On-line monitoring of process parameters during fermentation”. *Current Developments in Biotechnology and Bioengineering*. Szerk. Ranina Sirohi et al. United Kingdom: Elsevier, 117–164. old. DOI: [10.1016/B978-0-323-91167-2.00003-4](https://doi.org/10.1016/B978-0-323-91167-2.00003-4).
- Cavinato, Anna G., David M. Mayes, Zhihong. Ge  s James B. Callis (1990). “Noninvasive method for monitoring ethanol in fermentation processes using fiber-optic near-infrared spectroscopy”. *Analytical Chemistry* 62.18, 1977–1982. old. DOI: [10.1021/ac00217a015](https://doi.org/10.1021/ac00217a015).
- Ca ete-Carmona, Eduardo, Juan-Jos  Gallego-Mart nez, Cristian Mart n, Mar a Brox, Juan-Jes s Luna-Rodr guez  s Juan Moreno (2020). “A Low-Cost IoT Device to Monitor in Real-Time Wine Alcoholic Fermentation Evolution Through CO2 Emissions”. *IEEE Sensors Journal* 20.12, 6692–6700. old. DOI: [10.1109/JSEN.2020.2975284](https://doi.org/10.1109/JSEN.2020.2975284).
- Chapman, Pete, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer  s Rudiger Wirth (2000). *CRISP-DM 1.0 Step-by-step data mining guide*. Techn. jel. The CRISP-DM consortium.
- Chen, Lihua, Sanxia Wang, Bo Liu  s Shengbao Feng (2023). “Flavor characteristics of hullless barley wine fermented with mixed starters by molds and yeasts isolated from Jiuqu”. *Food Bioscience* 52, 102349. old. DOI: [10.1016/j.fbio.2022.102349](https://doi.org/10.1016/j.fbio.2022.102349).
- Cheng, Lu, Qing-Hao Meng, Achim J Lilienthal  s Pei-Feng Qi (2021). “Development of compact electronic noses: a review”. *Measurement Science and Technology* 32.6, 62002. old. DOI: [10.1088/1361-6501/abef3b](https://doi.org/10.1088/1361-6501/abef3b).
- Dymerski, Tomasz (2017). “Electronic Noses and Electronic Tongues”. *Advances in Biosensors: Reviews. Volume 1*. Szerk. Sergey Yurish. Barcelona, Spain: IFSA Publishing, S. L. 7. fej., 153–193. old.
- Enyedi, Fl rian, Huyen Trang Do Thi,  gnes Szanyi, P ter Mizsey, Andr s J zsef T th  s Tibor Nagy (2022). “Low-Cost and Efficient Solution for the Automation of Laboratory

- Scale Experiments: The Case of Distillation Column”. *Processes* 10.4, 737. old. DOI: [10.3390/pr10040737](https://doi.org/10.3390/pr10040737).
- Flach, Peter (2012). *Machine Learning. The Art and Science of Algorithms that Make Sense of Data*. New York, United States: Cambridge University Press. DOI: [10.1017/CB09780511973000](https://doi.org/10.1017/CB09780511973000).
- Ghasemi-Varnamkhasti, Mahdi, María Luz Rodríguez-Méndez, Seyed Saeid Mohtasebi, Constantin Apetrei, Jesus Lozano, Hojat Ahmadi, Seyed Hadi Razavi és Jose Antonio de Saja (2012). “Monitoring the aging of beers using a bioelectronic tongue”. *Food Control* 25.1, 216–224. old. DOI: [10.1016/j.foodcont.2011.10.020](https://doi.org/10.1016/j.foodcont.2011.10.020).
- Gil-Sánchez, Luis, Juan Soto, Ramón Martínez-Mañez, Eduardo Garcia-Breijo, Javier Ibáñez és Eduard Llobet (2011). “A novel humid electronic nose combined with an electronic tongue for assessing deterioration of wine”. *Sensors and Actuators A: Physical* 171.2, 152–158. old. DOI: [10.1016/j.sna.2011.08.006](https://doi.org/10.1016/j.sna.2011.08.006).
- Gillay, Zoltán (2023). *Méréstechnika*. Előadás anyagai (diasorok). Magyar Agrár- és Élettudományi Egyetem, Pálinkamester szaktanácsadó szakirányú továbbképzési szak.
- Gonzalez Viejo, Claudia, Sigfredo Fuentes, Amruta Godbole, Bryce Widdicombe és Ranjith R Unnithan (2020). “Development of a low-cost e-nose to assess aroma profiles: An artificial intelligence application to assess beer quality”. *Sensors and Actuators B: Chemical* 308, 127688. old. DOI: [10.1016/j.snb.2020.127688](https://doi.org/10.1016/j.snb.2020.127688).
- Gutiérrez, Juan Manuel, Zouhair Haddi, Aziz Amari, Benachir Bouchikhi, Aitor Mimen-dia, Xavier Cetó és Manel del Valle (2013). “Hybrid electronic tongue based on mul-tisensor data fusion for discrimination of beers”. *Sensors and Actuators B: Chemical* 177, 989–996. old. DOI: [10.1016/j.snb.2012.11.110](https://doi.org/10.1016/j.snb.2012.11.110).
- Hadi, Auday Shaker, Mohamed Alsaker, Ahmed Eshoom, Monaem Elmnifi, Mohammed A. Alhmode és Laith Jaafer Habeeb (2022). “Development of Low cost and Multi-Material Sensing Approach for MQ135 Sensor”. *ECS Transactions* 107.0, 75–89. old. DOI: [10.1149/10701.17309ecst](https://doi.org/10.1149/10701.17309ecst).
- Hawchar, Amer, Solomon Ould és Nick S. Bennett (2022). “Carbon Dioxide Monitoring inside an Australian Brewery Using an Internet-of-Things Sensor Network”. *Sensors* 22.24. DOI: [10.3390/s22249752](https://doi.org/10.3390/s22249752).
- He, Xi, Huang Yangming, Elżbieta Górská-Horczyzak, Agnieszka Wierzbicka és Henryk H. Jeleń (2021). “Rapid analysis of Baijiu volatile compounds fingerprint for their aroma and regional origin authenticity assessment”. *Food Chemistry* 337, 128002. old. DOI: [10.1016/j.foodchem.2020.128002](https://doi.org/10.1016/j.foodchem.2020.128002).
- Hoche, Sven, Daniel Krause, Mohamed A. Hussein és Thomas Becker (2016). “Ultrasound-based, in-line monitoring of anaerobe yeast fermentation: model, sensor design and process application”. *International Journal of Food Science & Technology* 51.3, 710–719. old. DOI: [10.1111/ijfs.13027](https://doi.org/10.1111/ijfs.13027).

- Hornik, Kurt, Maxwell Stinchcombe és Halbert White (1989). “Multilayer feedforward networks are universal approximators”. *Neural Networks* 2.5, 359–366. old. DOI: [10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- Horváth, Tomáš (2023a). *Introduction to Data Science*. Előadás anyagai (diasorok). Eötvös Loránd Tudományegyetem, Programtervező informatikus szak, Adattudományi specializáció.
- (2023b). *Machine Learning*. Előadás anyagai (diasorok). Eötvös Loránd Tudományegyetem, Programtervező informatikus szak, Adattudományi specializáció.
- Horváth, Tomáš és André C.P.L.F. de Carvalho (2017). “Evolutionary computing in recommender systems: a review of recent research”. *Natural Computing* 16, 441–462. old. DOI: [10.1007/s11047-016-9540-y](https://doi.org/10.1007/s11047-016-9540-y).
- Horváth, Tomáš, Rafael G. Mantovani és André C.P.L.F. de Carvalho (2023). “Hyperparameter initialization of classification algorithms using dynamic time warping: A perspective on PCA meta-features”. *Applied Soft Computing* 134, 109969. old. DOI: [10.1016/j.asoc.2022.109969](https://doi.org/10.1016/j.asoc.2022.109969).
- Hérics-Tóth, Jenő és Antal Osztróvszky (1918). *Szeszfőzés gyümölcsökből*. Budapest: Pátria Irodalmi Vállalat és Nyomdai Részvénytársaság.
- Jacques, K.A., T.P. Lyons és D.R. Kelsall (2003). *The Alcohol Textbook. A reference for the beverage, fuel and industrial alcohol industries*. 4. kiad. Nottingham, UK: Nottingham University Press.
- Karakaya, Diclehan, Oguzhan Ulucan és Mehmet Türkan (2020). “Electronic Nose and Its Applications: A Survey”. *International Journal of Automation and Computing* 17, 179–209. old. DOI: [10.1007/s11633-019-1212-9](https://doi.org/10.1007/s11633-019-1212-9).
- Kattipparambil Rajan, Dhanesh, Mimmi Patrikoski, Jarmo Verho, Jyrki Sivula, Heimo Ihalainen, Susanna Miettinen és Jukka Lekkala (2016). “Optical non-contact pH measurement in cell culture with sterilizable, modular parts”. *Talanta* 161, 755–761. old. DOI: [10.1016/j.talanta.2016.09.021](https://doi.org/10.1016/j.talanta.2016.09.021).
- Kovacs, Zoltan, Dániel Szöllősi, John-Lewis Zinia Zaukuu, Zsanett Bodor, Flóra Vitális, Balkis Aouadi, Viktória Zsom-Muha és Zoltan Gillay (2020). “Factors Influencing the Long-Term Stability of Electronic Tongue and Application of Improved Drift Correction Methods”. *Biosensors* 10.7, 74. old. DOI: [10.3390/bios10070074](https://doi.org/10.3390/bios10070074).
- Legin, A., A. Rudnitskaya, L. Lvova, Yu. Vlasov, C. Di Natale és A. D’Amico (2003). “Evaluation of Italian wine by the electronic tongue: recognition, quantitative analysis and correlation with human sensory perception”. *Analytica Chimica Acta* 484.1, 33–44. old. DOI: [10.1016/S0003-2670\(03\)00301-5](https://doi.org/10.1016/S0003-2670(03)00301-5).
- Li, Bingyang és Yu Gu (2023). “A Machine Learning Method for the Quality Detection of Base Liquor and Commercial Liquor Using Multidimensional Signals from an Electronic Nose”. *Foods* 12.7. DOI: [10.3390/foods12071508](https://doi.org/10.3390/foods12071508).

- Lidgren, Lina, Ola Lilja, Margareta Krook és Dario Kriz (2006). “Automatic fermentation control based on a real-time in situ SIRE® biosensor regulated glucose feed”. *Biosensors and Bioelectronics* 21.10, 2010–2013. old. DOI: [10.1016/j.bios.2005.09.012](https://doi.org/10.1016/j.bios.2005.09.012).
- Lidén, Helena, Thomas Bachinger, Lo Gorton és Carl-Fredrik Mandenius (2000). “On-line determination of non-volatile or low-concentration metabolites in a yeast cultivation using an electronic nose”. *Analyst* 125 (6), 1123–1128. old. DOI: [10.1039/A909585C](https://doi.org/10.1039/A909585C).
- Lim, Jong Hyun és Tai Hyun Park (2014). “Concept of Bioelectronic Nose”. *Bioelectronic Nose: Integration of Biotechnology and Nanotechnology*. Szerk. Tai Hyun Park. Dordrecht: Springer Netherlands, 1–22. old. DOI: [10.1007/978-94-017-8613-3_1](https://doi.org/10.1007/978-94-017-8613-3_1).
- Lin, Jessica, Eamonn Keogh, Stefano Lonardi és Pranav Patel (2002). “Finding Motifs in Time Series”. *Proceedings of the Second Workshop on Temporal Data Mining*.
- Lin, Jessica, Eamonn Keogh, Li Wei és Stefano Lonardi (2007). “Experiencing SAX: a novel symbolic representation of time series”. *Data Mining and Knowledge Discovery* 15.2, 107–144. old. DOI: [10.1007/s10618-007-0064-z](https://doi.org/10.1007/s10618-007-0064-z).
- Liu, Huixiang, Qing Li, Bin Yan, Lei Zhang és Yu Gu (2019). “Bionic Electronic Nose Based on MOS Sensors Array and Machine Learning Algorithms Used for Wine Properties Detection”. *Sensors* 19.1. DOI: [10.3390/s19010045](https://doi.org/10.3390/s19010045).
- Liu, Taoping, Lihua Guo, Mou Wang, Chen Su, Di Wang, Hao Dong, Jingdong Chen és Weiwei Wu (2023). “Review on Algorithm Design in Electronic Noses: Challenges, Status, and Trends”. *Intelligent Computing* 2, 12. old. DOI: [10.34133/icomputing.0012](https://doi.org/10.34133/icomputing.0012).
- Lorwongtragool, Panida, Chatchawal Wongchoosuk és Teerakiat Kerdcharoen (2011). “Portable electronic nose for beverage quality assessment”. *The 8th ECTI Association of Thailand – Conference 2011*, 163–166. old. DOI: [10.1109/ECTICON.2011.5947797](https://doi.org/10.1109/ECTICON.2011.5947797).
- Lozano, Jesús, José Pedro Santos és M. Carmen Horrillo (2016). “Chapter 14 - Wine Applications With Electronic Noses”. *Electronic Noses and Tongues in Food Science*. Szerk. María Luz Rodríguez Méndez. San Diego: Academic Press, 137–148. old. DOI: [10.1016/B978-0-12-800243-8.00014-7](https://doi.org/10.1016/B978-0-12-800243-8.00014-7).
- Macías, Miguel Macías, J. Enrique Agudo, Antonio García Manso, Carlos Javier García Orellana, Horacio Manuel González Velasco és Ramón Gallardo Caballero (2013a). “A Compact and Low Cost Electronic Nose for Aroma Detection”. *Sensors* 13.5, 5528–5541. old. DOI: [10.3390/s130505528](https://doi.org/10.3390/s130505528).
- Macías, Miguel Macías, Antonio García Manso, Carlos Javier García Orellana, Horacio Manuel González Velasco, Ramón Gallardo Caballero és Juan Carlos Peguero Chamizo (2013b). “Acetic Acid Detection Threshold in Synthetic Wine Samples of a Portable Electronic Nose”. *Sensors* 13.1, 208–220. old. DOI: [10.3390/s130100208](https://doi.org/10.3390/s130100208).
- Mamolar-Domenech, S., H. Crespo-Sariol, J.C. Sáenz-Díez, A. Sánchez-Roca, Juan-Ignacio Latorre-Biel és J. Blanco (2023). “A new approach for monitoring the alcoholic ferment-

- tation process based on acoustic emission analysis: A preliminary assessment”. *Journal of Food Engineering* 353, 111537. old. DOI: [10.1016/j.jfoodeng.2023.111537](https://doi.org/10.1016/j.jfoodeng.2023.111537).
- Mendez, M.L.R. és V.R. Preedy (2016). *Electronic Noses and Tongues in Food Science*. Elsevier Science. DOI: [10.1016/C2013-0-14449-2](https://doi.org/10.1016/C2013-0-14449-2).
- Montague, G.A., A.J. Morris és A.C. Ward (1989). “Fermentation Monitoring and Control: A Perspective”. *Biotechnology and Genetic Engineering Reviews* 7.1, 147–188. old. DOI: [10.1080/02648725.1989.10647858](https://doi.org/10.1080/02648725.1989.10647858).
- Moreira, João, André C.P.L.F. de Carvalho és Tomáš Horváth (2018). *A General Introduction to Data Analytics*. Hoboken, USA: John Wiley & Sons. DOI: [10.1002/9781119296294](https://doi.org/10.1002/9781119296294).
- Nery, Emilia Witkowska és Lauro T. Kubota (2016). “Integrated, paper-based potentiometric electronic tongue for the analysis of beer and wine”. *Analytica Chimica Acta* 918, 60–68. old. DOI: [10.1016/j.aca.2016.03.004](https://doi.org/10.1016/j.aca.2016.03.004).
- Nyitrai, Ákos, Attila Gere és László Sipos (2018). “Mesterséges neurális hálózatok élelmiszertudományi alkalmazásai és nemzetközi trendjei”. *Élelmiszervizsgálati közlemények* 64.3.
- Offner, Jim (2018). “Automation Now Making a Splash in Distilleries”. *Beverage Master: Craft Spirits & Brew Magazine* December 2017 – January 2018, 11–19. old.
- Paknahad, Mohammad, Ali Ahmadi, Jacques Rousseau, Hojatollah Rezaei Nejad és Mina Hoorfar (2017). “On-Chip Electronic Nose For Wine Tasting: A Digital Microfluidic Approach”. *IEEE Sensors Journal* 17.14, 4322–4329. old. DOI: [10.1109/JSEN.2017.2707525](https://doi.org/10.1109/JSEN.2017.2707525).
- Panyik, Gáborné (2013). *Pálinkafőzés. Ágyas pálinka és likőr készítése*. Budapest, Hungary: Cser Kiadó.
- Peris, Miguel és Laura Escuder Gilabert (2013). “On-line monitoring of food fermentation processes using electronic noses and electronic tongues: A review”. *Analytica Chimica Acta* 804, 29–36. old. DOI: [10.1016/j.aca.2013.09.048](https://doi.org/10.1016/j.aca.2013.09.048).
- Pinheiro, Carmen, Carla M. Rodrigues, Thomas Schäfer és João G. Crespo (2002). “Monitoring the aroma production during wine–must fermentation with an electronic nose”. *Biotechnology and Bioengineering* 77.6, 632–640. old. DOI: [10.1002/bit.10141](https://doi.org/10.1002/bit.10141).
- Qi, Pei-Feng, Qing-Hao Meng, Ya-Qi Jing, Ying-Jie Liu és Ming Zeng (2017a). “A Bio-Inspired Breathing Sampling Electronic Nose for Rapid Detection of Chinese Liquors”. *IEEE Sensors Journal* 17.15, 4689–4698. old. DOI: [10.1109/JSEN.2017.2712742](https://doi.org/10.1109/JSEN.2017.2712742).
- Qi, Pei-Feng, Qing-Hao Meng és Ming Zeng (2017b). “A CNN-based simplified data processing method for electronic noses”. *ISOCs/IEEE International Symposium on Olfaction and Electronic Nose*, 1–3. old. DOI: [10.1109/ISOEN.2017.7968887](https://doi.org/10.1109/ISOEN.2017.7968887).
- Qi, Pei-Feng, Qing-Hao Meng, Yu Zhou, Ya-Qi Jing és Ming Zeng (2015). “A portable E-nose system for classification of Chinese liquor”. *2015 IEEE SENSORS*, 1–4. old. DOI: [10.1109/ICSENS.2015.7370248](https://doi.org/10.1109/ICSENS.2015.7370248).

- Qi, Pei-Feng, Ming Zeng, Zhi-Hua Li, Biao Sun és Qing-Hao Meng (2017c). “Design of a portable electronic nose for real-fake detection of liquors”. *Review of Scientific Instruments* 88.9, 95001. old. DOI: [10.1063/1.5001314](https://doi.org/10.1063/1.5001314).
- Ragazzo-Sanchez, J.A., P. Chalier, D. Chevalier-Lucia, M. Calderon-Santoyo és C. Ghommidh (2009). “Off-flavours detection in alcoholic beverages by electronic nose coupled to GC”. *Sensors and Actuators B: Chemical* 140.1, 29–34. old. DOI: [10.1016/j.snb.2009.02.061](https://doi.org/10.1016/j.snb.2009.02.061).
- Rahmatnejad, Vida, Yunqian Wei és Govind Rao (2023). “Recent Developments in Bioprocess Monitoring Systems”. *Biopharmaceutical Manufacturing: Progress, Trends and Challenges*. Szerk. Ralf Pörtner. Cham: Springer International Publishing, 39–66. old. DOI: [10.1007/978-3-031-45669-5_2](https://doi.org/10.1007/978-3-031-45669-5_2).
- Ranasinghe, Damith C., Nickolas J. G. Falkner, Pan Chao és Wu Hao (2013). “Wireless sensing platform for remote monitoring and control of wine fermentation”. *2013 IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, 503–508. old. DOI: [10.1109/ISSNIP.2013.6529841](https://doi.org/10.1109/ISSNIP.2013.6529841).
- Ratanamahatana, Chotirat Ann és Eamonn Keogh (2004). “Everything you know about dynamic time warping is wrong”. *Third Workshop on Mining Temporal and Sequential Data*. Citeseer.
- Rodriguez-Mendez, M.L., A.A. Arrieta, V. Parra, A. Bernal, A. Vegas, S. Villanueva, R. Gutierrez-Osuna és J.A. de Saja (2004). “Fusion of three sensory modalities for the multimodal characterization of red wines”. *IEEE Sensors Journal* 4.3, 348–354. old. DOI: [10.1109/JSEN.2004.824236](https://doi.org/10.1109/JSEN.2004.824236).
- Rodríguez-Méndez, María, J.A. De Saja, Rocio Gonzalez-Anton, Celia García-Hernández, Cristina Medina Plaza, Cristina García-Cabezón és Fernando Martin (2016). “Electronic Noses and Tongues in Wine Industry”. *Frontiers in Bioengineering and Biotechnology* 4. DOI: [10.3389/fbioe.2016.00081](https://doi.org/10.3389/fbioe.2016.00081).
- Rodríguez Méndez, María Luz, José A. De Saja, C. Medina-Plaza és C. García-Hernández (2016). “Chapter 26 - Electronic Tongues for the Organoleptic Characterization of Wines”. *Electronic Noses and Tongues in Food Science*. Szerk. María Luz Rodríguez Méndez. San Diego: Academic Press, 265–273. old. DOI: [10.1016/B978-0-12-800243-8.00026-3](https://doi.org/10.1016/B978-0-12-800243-8.00026-3).
- Sainz, Beatriz, Jonathan Antolín, Miguel López-Coronado és Carlos De Castro (2013). “A Novel Low-Cost Sensor Prototype for Monitoring Temperature during Wine Fermentation in Tanks”. *Sensors* 13.3, 2848–2861. old. DOI: [10.3390/s130302848](https://doi.org/10.3390/s130302848).
- Santos, J. P., J. M. Cabellos, T. Arroyo és M. C. Horrillo (2011). “Portable Electronic Nose to Discriminate Artificial Aged Wine from Barrel-Aged Wine”. *AIP Conference Proceedings* 1362.1, 171–173. old. DOI: [10.1063/1.3626347](https://doi.org/10.1063/1.3626347).

- Santos, José Pedro, Jesús Lozano és Manuel Aleixandre (2017). “Electronic Noses Applications in Beer Technology”. *Brewing Technology*. Szerk. Makoto Kanauchi. Rijeka: IntechOpen. 9. fejr. DOI: [10.5772/intechopen.68822](https://doi.org/10.5772/intechopen.68822).
- Scrimgeour, Neil (2000). *Evaluating the viability of process sensor technologies for measurement of sugar levels during fermentation. Final report to Australian Grape and Wine Authority (Project Number: AWR 1401)*. Techn. jel. Australian Wine Research Institute.
- Seesaard, Thara és Chatchawal Wongchoosuk (2022). “Recent Progress in Electronic Noses for Fermented Foods and Beverages Applications”. *Fermentation* 8.7. DOI: [10.3390/fermentation8070302](https://doi.org/10.3390/fermentation8070302).
- Seixas Jr., José Luis és Tomáš Horváth (2020. szept.). “KNN algorithm with DTW distance for signature classification of wine leaves”. *20th Conference Information Technologies - Applications and Theory (ITAT 2020)*. Oravská Lesná, Slovakia, 130–136. old.
- Shi, Yan, Furong Gong, Mingyang Wang, Jingjing Liu, Yinong Wu és Hong Men (2019). “A deep feature mining method of electronic nose sensor data for identifying beer olfactory information”. *Journal of Food Engineering* 263, 437–445. old. DOI: [10.1016/j.jfoodeng.2019.07.023](https://doi.org/10.1016/j.jfoodeng.2019.07.023).
- Shi, Zhi-biao, Tao Yu, Qun Zhao, Yang Li és Yu-bin Lan (2008). “Comparison of Algorithms for an Electronic Nose in Identifying Liquors”. *Journal of Bionic Engineering* 5.3, 253–257. old. DOI: [10.1016/S1672-6529\(08\)60032-3](https://doi.org/10.1016/S1672-6529(08)60032-3).
- Shrake, Nicholas L., Rajeevan Amirtharajah, Charles Brenneman, Roger Boulton és André Knoesen (2014). “In-Line Measurement of Color and Total Phenolics during Red Wine Fermentations Using a Light-Emitting Diode Sensor”. *American Journal of Enology and Viticulture* 65.4, 463–470. old. DOI: [10.5344/ajev.2014.14023](https://doi.org/10.5344/ajev.2014.14023).
- Srinivasan, K., Prem B. Madheswaran, Kumar M. Sathish és A. Thirupathi (2013). “Automation in Rectified Spirit (RS) Production in Distillery Unit”. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* 2.2, 909–916. old.
- Tamo, Alberth és Alexander Hilario-Tacuri (2020). “Implementing WLAN-IoT control system for brewing fermentation through Raspberry PI”. *IEEE International Conference on Electronics, Electrical Engineering and Computing*. ACM, 1–4. old. DOI: [10.1109/INTERCON50315.2020.9220232](https://doi.org/10.1109/INTERCON50315.2020.9220232).
- Tan, Juzhong és Jie Xu (2020). “Applications of electronic nose (e-nose) and electronic tongue (e-tongue) in food quality-related properties determination: A review”. *Artificial Intelligence in Agriculture* 4, 104–115. old. DOI: [10.1016/j.aiia.2020.06.003](https://doi.org/10.1016/j.aiia.2020.06.003).
- Toko, Kiyoshi és Yusuke Tahara (2016). “Chapter 16 - Beer Analysis Using an Electronic Tongue”. *Electronic Noses and Tongues in Food Science*. Szerk. María Luz Rodríguez Méndez. San Diego: Academic Press, 161–170. old. DOI: [10.1016/B978-0-12-800243-8.00016-0](https://doi.org/10.1016/B978-0-12-800243-8.00016-0).

- Tomtsis, Dimitrios, Sotirios Kontogiannis, George Kokkonis és Nicholas Zinas (2016). “IoT Architecture for Monitoring Wine Fermentation Process of Debina Variety Semi-Sparkling Wine”. *SouthEast European Design Automation, Computer Engineering, Computer Networks and Social Media Conference*. New York, USA: Association for Computing Machinery, 42–47. DOI: [10.1145/2984393.2984398](https://doi.org/10.1145/2984393.2984398).
- Toyota, Kentaro, Hong Cui, Kentaro Abe, Masaaki Habara, Kiyoshi Toko és Hidekazu Ikezaki (2011). “Sweetness sensor with lipid/polymer membranes: Response to various sugars”. *Sensors and Materials* 23.8, 475–482. old. DOI: [10.18494/SAM.2011.772](https://doi.org/10.18494/SAM.2011.772).
- Uehara, Yui, Satoshi Ohtake és Takamoto Fukura (2018). “A Mash Temperature Monitoring System for Sake Brewing”. *IEEE International Conference on Consumer Electronics-Taiwan*. ACM, 1–2. old. DOI: [10.1109/ICCE-China.2018.8448774](https://doi.org/10.1109/ICCE-China.2018.8448774).
- Uehara, Yui, Satoshi Ohtake és Takayuki Karyu (2019). “A Koji Temperature Monitoring System”. *IEEE International Conference on Consumer Electronics - Taiwan*, 1–2. old. DOI: [10.1109/ICCE-TW46550.2019.8991866](https://doi.org/10.1109/ICCE-TW46550.2019.8991866).
- Ujszászi, János és Lajos Sólyom (1980). *Szeszesítel-ipari vizsgálati módszerek*. Budapest, Hungary: Mezőgazdasági Könyvkiadó.
- Varnamkhasti, M. Ghasemi, S.S. Mohtasebi, M.L. Rodriguez-Mendez, J. Lozano, S.H. Razavi és H. Ahmadi (2011). “Potential application of electronic nose technology in brewery”. *Trends in Food Science & Technology* 22.4, 165–174. old. DOI: [10.1016/j.tifs.2010.12.005](https://doi.org/10.1016/j.tifs.2010.12.005).
- Vasilescu, Alina, Pablo Fanjul-Bolado, Ana-Maria Titoiu, Roxana Porumb és Petru Epure (2019). “Progress in Electrochemical (Bio)Sensors for Monitoring Wine Production”. *Chemosensors* 7.4. DOI: [10.3390/chemosensors7040066](https://doi.org/10.3390/chemosensors7040066).
- Violino, Simona, Simone Figorilli, Corrado Costa és Federico Pallottino (2020). “Internet of Beer: A Review on Smart Technologies from Mash to Pint”. *Foods* 9.7. DOI: [10.3390/foods9070950](https://doi.org/10.3390/foods9070950).
- Vivaldi, Federico, Pietro Salvo, Noemi Poma, Andrea Bonini, Denise Biagini, Lorenzo Del Noce, Bernardo Melai, Fabio Lisi és Fabio Di Francesco (2021). “Recent Advances in Optical, Electrochemical and Field Effect pH Sensors”. *Chemosensors* 9.2. DOI: [10.3390/chemosensors9020033](https://doi.org/10.3390/chemosensors9020033).
- Vošahlík, Jan és J. Hart (2021). “Measurability of quality in fermentation process of rice wine by IoT in the field of industry 4.0”. *Agronomy Research* 19.S3, 1318–1324. DOI: [10.15159/AR.21.063](https://doi.org/10.15159/AR.21.063).
- Wang, Jia-Ying, Qing-Hao Meng, Xing-Wei Jin és Zhe-Hua Sun (2021). “Design of hand-held electronic nose bionic chambers for Chinese liquors recognition”. *Measurement* 172, 108856. old. DOI: [10.1016/j.measurement.2020.108856](https://doi.org/10.1016/j.measurement.2020.108856).
- Wang, Mingyang és Yinsheng Chen (2024). “Electronic nose and its application in the food industry: a review”. *European Food Research and Technology* 250, 21–67. DOI: [10.1007/s00217-023-04381-z](https://doi.org/10.1007/s00217-023-04381-z).

- Wang, Song, Xiao-Zhen Hu, Yan-Yan Liu, Ning-Ping Tao, Ying Lu, Xi-Chang Wang, Wing Lam, Ling Lin és Chang-Hua Xu (2022). “Direct authentication and composition quantitation of red wines based on Tri-step infrared spectroscopy and multivariate data fusion”. *Food Chemistry* 372, 131259. old. DOI: [10.1016/j.foodchem.2021.131259](https://doi.org/10.1016/j.foodchem.2021.131259).
- Wolpert, David H. (1996). “The Lack of A Priori Distinctions Between Learning Algorithms”. *Neural Computation* 8.7, 1341–1390. old. DOI: [10.1162/neco.1996.8.7.1341](https://doi.org/10.1162/neco.1996.8.7.1341).
- Yang, Yan, Huixiang Liu és Yu Gu (2020). “A Model Transfer Learning Framework With Back-Propagation Neural Network for Wine and Chinese Liquor Detection by Electronic Nose”. *IEEE Access* 8, 105278–105285. old. DOI: [10.1109/ACCESS.2020.2999591](https://doi.org/10.1109/ACCESS.2020.2999591).
- Ye, Lexiang és Eamonn Keogh (2009). “Time Series Shapelets: A New Primitive for Data Mining”. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, USA: Association for Computing Machinery, 947–956. DOI: [10.1145/1557019.1557122](https://doi.org/10.1145/1557019.1557122).
- Ye, Zhenyi, Yuan Liu és Qiliang Li (2021). “Recent Progress in Smart Electronic Nose Technologies Enabled with Machine Learning Methods”. *Sensors* 21, 7620. old. DOI: [10.3390/s21227620](https://doi.org/10.3390/s21227620).
- Zhang, Shanshan, Yu Cheng, Dehan Luo, Jiafeng He, Angus K. Y. Wong és Kevin Hung (2021). “Channel Attention Convolutional Neural Network for Chinese Baijiu Detection With E-Nose”. *IEEE Sensors Journal* 21.14, 16170–16182. old. DOI: [10.1109/JSEN.2021.3075703](https://doi.org/10.1109/JSEN.2021.3075703).
- Zhang, Shengyi és Xinming Wang (2013). “The Design of Temperature Control System for Rice Wine Fermentation”. *International Conference On Systems Engineering and Modeling*. Atlantis Press, 1007–1010. old. DOI: [10.2991/icsem.2013.212](https://doi.org/10.2991/icsem.2013.212).
- Zhang, Wentian, Taoping Liu, Amber Brown, Maiken Ueland, Shari L. Forbes és Steven Weidong Su (2022). “The Use of Electronic Nose for the Classification of Blended and Single Malt Scotch Whisky”. *IEEE Sensors Journal* 22.7, 7015–7021. old. DOI: [10.1109/JSEN.2022.3147185](https://doi.org/10.1109/JSEN.2022.3147185).
- Zhao, Wei, Qing-Hao Meng, Ming Zeng és Pei-Feng Qi (2017). “Stacked Sparse Auto-Encoders (SSAE) Based Electronic Nose for Chinese Liquors Classification”. *Sensors* 17.12. DOI: [10.3390/s17122855](https://doi.org/10.3390/s17122855).
- Zhu, Xianglin, Khalil Ur Rehman, Bo Wang és Muhammad Shahzad (2020). “Modern Soft-Sensing Modeling Methods for Fermentation Processes”. *Sensors* 20.6, 1771. old. DOI: [10.3390/s20061771](https://doi.org/10.3390/s20061771).

Ábrák jegyzéke

2.1. Az IE monitorozásának módjai (forrás: Cargalo és tsai. (2022)).	17
5.1. A javasolt erjedést monitorozó rendszer leegyszerűsített vázlata	38

Táblázatok jegyzéke

2.1. MQ típusú szenzorok listája (https://robu.in/mq-series-gas-sensor/).	5
2.2. A főbb adatbányászati és gépi tanulás módszerek rendszerezése	8
2.3. Survey cikkek szenzorok élelmiszeripari alkalmazásairól a bennük található összes, a szeszesitalokkal kapcsolatos és az alkoholos erjesztéssel kapcsolatos idézett cikkek számával („eO” = e-Orr, „eN” = e-Nyelv, „eSz” = egyéb szenzor).	22
2.4. Alkoholos erjesztés in-line és on-line monitorozását tárgyaló cikkek, jelezve a bennük bemutatott kutatás reprodukálhatóságát („R” oszlop) és a bemutatott módszerek megvalósíthatóságát („M” oszlop). Jelölések: „eO” = e-Orr, „eN” = e-Nyelv, „oSz” = optikai szenzor, „GC” = gázkromatográf, „eSz” = egyéb szenzor.	28

NYILATKOZAT

a szakdolgozat nyilvános hozzáféréséről és eredetiségéről

A hallgató neve: Horváth Tomás
A Hallgató Neptun kódja: S9KPG7
A dolgozat címe: Szenzorok alkalmazási lehetőségei az erjedés monitorozásához
A megjelenés éve: 2024
A konzulens intézetének neve: Élelmiszertudományi és Technológiai Intézet
A konzulens tanszékének a neve: Biomérnök és Erjedésipari Technológia Tanszék,
Árukezelési, Kereskedelmi, Ellátási Lánc és
Érzékszervi Minősítési Tanszék

Kijelentem, hogy az általam benyújtott szakdolgozat egyéni, eredeti jellegű, saját szellemi alkotásom. Azon részeket, melyeket más szerzők munkájából vettem át, egyértelműen megjelöltem, és az irodalomjegyzékben szerepeltettem.

Ha a fenti nyilatkozattal valótlan állítottam, tudomásul veszem, hogy a záróvizsga-bizottság a záróvizsgából kizár és a záróvizsgát csak új dolgozat készítése után tehetek.

A leadott dolgozat, mely PDF dokumentum, szerkesztését nem, megtekintését és nyomtatását engedélyezem.

Tudomásul veszem, hogy az általam készített dolgozatra, mint szellemi alkotás felhasználására, hasznosítására a Magyar Agrár- és Élettudományi Egyetem mindenkori szellemitulajdon-kezelési szabályzatában megfogalmazottak érvényesek.

Tudomásul veszem, hogy dolgozatom elektronikus változata feltöltésre kerül a Magyar Agrár- és Élettudományi Egyetem könyvtári repozitori rendszerébe. Tudomásul veszem, hogy a megvédett és

- nem titkosított dolgozat a védelmet követően
- titkosításra engedélyezett dolgozat a benyújtásától számított 5 év eltelté után nyilvánosan elérhető és kereshető lesz az Egyetem könyvtári repozitori rendszerében.

Kelt: Edinburgh, Skócia 2024. év április hó 26. nap



Hallgató aláírása

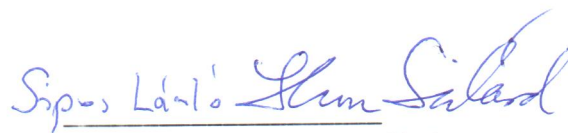
NYILATKOZAT

Horváth Tomás (hallgató Neptun azonosítója: S9KPG7) konzulenseként nyilatkozom arról, hogy szakdolgozatot áttekintettem, a hallgatót az irodalmi források korrekt kezelésének követelményeiről, jogi és etikai szabályairól tájékoztattam.

A szakdolgozatot a záróvizsgán történő védeésre javaslom / nem javaslom¹.

A dolgozat állam- vagy szolgálati titkot tartalmaz: igen nem^{*2}

Kelt: Budapest, 2024 év április hó 20. nap



Dr. Sipos László és Dr. Kun Szilárd
konzulensek

¹ A megfelelő aláhúzendó.

² A megfelelő aláhúzendó.