



A kép illusztráció / Picture is for illustration only
Fotó/Photo: Pixabay

Sipos László¹, Losó Viktor¹, Nyitrai Ákos¹, Kókai Zoltán¹, Gere Attila¹

Érkezett: 2017. július – Elfogadva: 2017. október

Érzékszervi kedveltség predikciója mesterséges neurális hálózatokkal, fagyasztott csemegekukorica-fajták példáján bemutatva

Kulcsszavak: predikció, fogyasztói kedveltség, többrétegű előreccsatolt neurális háló (MLFN), csemegekukorica, csemegekukorica beltartalmi jellemzői, klaszteranalízis

1. Összefoglalás

A nemzeti és nemzetközi fajtajegyzékben található csemegekukorica (*Zea mays* var. *saccharata* L.) hibridek fajtákra lebontott érzékszervi profiljellemzőiről, fogyasztói preferenciáiról a rendelkezésünkre álló irodalmi forrásokban ismereteink szerint csak néhány publikáció született. Kutatásunkban a mesterséges neurális hálózatok (artificial neural networks, ANNs) gyakorlati alkalmazását mutatjuk be. Vizsgálatunkban 41 fagyasztott csemegekukorica-fajtát egy szakértői érzékszervi bírálócsoporthoz értékelt (14 fő), teljeskörű profilanalízis módszerével (MSZ ISO 11035:2001; ISO 13299:2003), 0-100-ig terjedő strukturálatlan skálán, majd nagymintás tesztben, fogyasztók (167 fő) a 41 fajta közül 6 fajtát jellemeztek kedveltség alapján, 9 elemű strukturált skálán.

A mesterséges neurális hálózatok nagymennyiségű adatot igényelnek, ezért a 6 fajta elkészült szakértői és fogyasztói adatokon 1000-szeres Monte Carlo szimulációt futtattunk, amelynek 80 %-án tréningeztük, 20 %-án pedig teszteltük a létrejött neurális hálókat. A legjobb predikciót a 4 nódusos többrétegű előreccsatolt (multi-layer feedforward neural net, MLFN) adta, ebben az esetben adódtak a legkisebb maradékok a tréning és a teszt során, amelyeket véletlen számokon történő előrejelzéssel, és keresztellenőrzéssel is validáltunk. Ezzel a felépített modellel jeleztük előre a többi 35 kukoricafajta kedveltségi értékét. A leginkább kedveltnek a 'Shinerock' fajta (8,46), míg a predikciók szerint a legkevésbé kedvelt a 'Madonna' és a 'Rustler' fajták lettek 2,7-es átlagos kedveltségi értékekkel rendelkeztek (1-9 tagú skálán).

A mesterséges neurális hálózat modell megalkotása során sikeresen azonosítottuk azokat a terméktulajdonságokat is, amelyek a fogyasztói elfogadás fő mozgatórugói: édes íz, globális ízintenzitás és lédúság. Összefoglalóan megállapítható, hogy a bemutatott validált termékspecifikus mesterséges neurális hálózat lehetővé teszi az egyes fajtákra vonatkoztatott kedveltség előrejelzését.

2. Bevezetés és szakirodalmi áttekintés

A mesterséges neurális hálózatok kialakulásához kulcsfontosságú volt az analógia, az emberi idegrendszer felépítésének és működésének feltárása. Az ideghálózat-programokat eredetileg az idegrendszer modelljeként fejlesztették, ahol a bemenetek összegyűjtik a beérkező jeleket más neuronoktól, majd a feldolgozó egység (neuron) elvégzi az összegzést, ez-

után az eredménytől függően a kimenetek továbbítják a jelet [1], [2], [3], [4], [5]. A mesterséges ideghálózatok kutatásában áttörést jelentettek Hopfield [6], Rumelhart és munkatársai [7] kutatásai, amelyekben az ideghálózati programok dinamikus modellezésével megoldották a nemlineáris leképezést, valamint az outputok inputokkal történő visszacsatolását. A mai értelemben vett mesterséges neurális hálózatok olyan párhuzamos működésre képes egyszerű fel-

¹ Szent István Egyetem, Élelmiszertudományi Kar, Árukezelési és Érzékszervi Minősítési Tanszék

dolgozó alapegységek (neuron) összekapcsolásából felépülő hálózatba szervezett rendszerek, amelyek tanulási és információ előhívási algoritmusokkal rendelkeznek.

A mesterséges ideghálózatok leegyszerűsítve olyan rendszerek, amelyek tanulnak a múltbeli tapasztalatokból, és a megtanultakat képesek előhívni. Legfőbb előnyeik a non-linearitás, a jó illeszkedés, a párhuzamos számítások, a gyors számítási sebesség és az adaptálhatóság. A legfontosabb felhasználási területei a kiugró értékek azonosítása, a korreláció azonosítása változók között, térredukció, regresszió nem lineáris változók között, komplex kapcsolatok modellezése, osztályozás, kategóriába sorolás [8], [9].

A neurális hálózatok legfontosabb tényezői a hálózatot felépítő elemek (neuronok/csomópontok/nóduszok), a hálózat összeköttetései (szerkezeti felépítés/topológia) és a tanulási algoritmusok. Az ANN működési elve a funkcionális kapcsolat modellezésén alapszik a bemenő (input) és a megfelelő kimeneti (output) változók között: $y = f(x)$, ahol x és y az input és az output vektorok, illetve, hogy az f szimbólum mutatja, hogy funkcionális kapcsolat van. A neuronokba beérkező jel a hozzátartozó súlytényezővel szorozódik, majd ezek összegződnek. Ezután a neuron – jellemzően nemlineáris átalakító függvény szerint – lépcsős, tanges hiperbolikus, logisztikus, szigmoid, stb. – kiszámítja a kimeneti jel értékét (output). Ezek a neuron egységek kapcsolódhatnak további rétegek neuronjaihoz, egymásnak adják át az információt a végső kimenetig, ahol a bemeneti jel kimeneti jellé alakul [10].

A rétegek és a neuronok számának meghatározása különösen fontos, mivel ez dönti el, hogy a hálózat képes lesz-e megtanulni a független és a függő változók közti kapcsolatokat. A rejtett rétegek száma és a rejtett rétegek neuronjainak száma az osztályozási feladat komplexitásától és az adatok mennyiségétől függ. Általánosságban az egy rejtett réteget és két rejtett neuront tartalmazó hálózatok nem tréningezhetőek kielégítő hibaszintre. Egy rejtett réteg és szigmoid aktivációs függvény segítségével bármilyen lineáris és nemlineáris probléma megoldható. Két vagy több rejtett réteget használata szükség nélkül megnöveli a tréning idejét. Egy rejtett rétegben általában már néhány neuron is elegendő. Az input réteg elágazásainak száma (neuronok) összhangban van az osztályozott objektumokat leíró változók számával, míg az output réteg elágazásainak száma az osztályok számával [11], [12], [13].

A hálózatok tanulási módja alapján megkülönböztetjük a felügyelt (*supervised*) és felügyeletlen (*unsupervised*) hálózatokat. A felügyelt tanulási hálózatokat összeköttetési, szerkezeti rendszere (topológiája) szerint pedig előrecsatolt hálózatok (*feedforward networks*) és hurkot tartalmazó visszacsatolt hálózatok (*recurrent networks*) lehetnek. A többrétegű előrecsatolt hálózatokban az információ előre áramlik úgy, hogy az adott rétegben levő neuronok nem kapnak jelet addig, amíg az előző réteg egységei azt elő nem állították, így minden réteg output vektora a következő réteg input vektora is egyben. A nem felügyelt tanulási hálózatokhoz tartoznak a Kohonen alapú szerkezetek [10]. A neurális hálózatok gyakorlatban történő alkalmazása három fő lépésből áll. A tanulási

fázis során épül fel a neurális háló, a validáló lépésben történik a felépített neurális háló modellmutatókkal történő érvényesítése, illetve a tesztelő lépéssel vizsgálható a neurális háló használhatósága [5], [11].

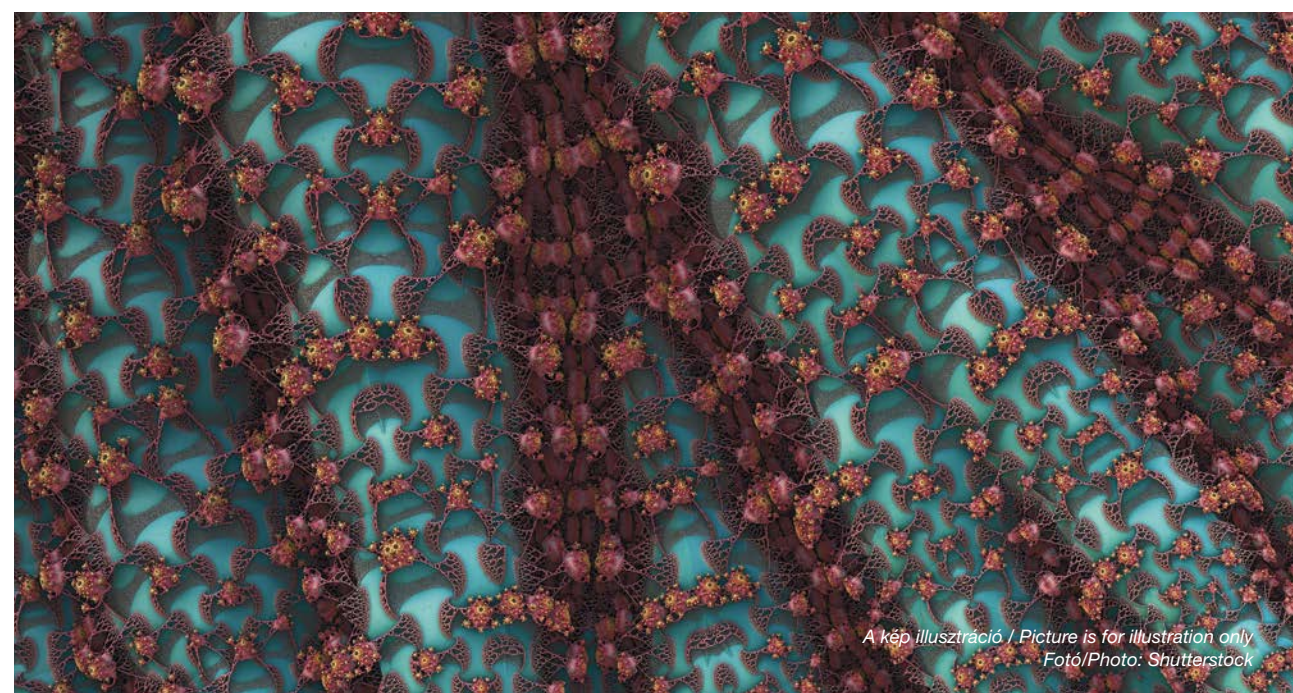
A továbbiakban a vizsgálatunk tárgyát képező csemegekukoricával kapcsolatos fontosabb dietetikai, termesztési, nemesítési összefüggéseket mutatjuk be. A zöldségfélék nagymértékben hozzájárulnak az emberi szervezetben a vitamin- és ásványi anyag fedezéséhez, befolyásolják a szénhidrát- és zsírszanyagcserét és a szervezetben lejátszódó összes olyan folyamat működését, amely rosttartalmukkal áll kapcsolatban [14]. A csemegekukorica (*Zea mays var saccharata* L.) az átlagos zöldségekhez képest sok energiát tartalmaz. Elsősorban a szénhidrát- és fehérjetartama adja a táplálkozási jelentőségét. A csemegekukoricák fajták szénhidrátprofiljainak jellemzője, hogy a fajtákban átlagosan a glükózból és fruktózból létrejött szacharóz (diszacharid) relatív súlya a legmagasabb (85%), utána következik a glükóz (10%), majd a fruktózhhoz (5%). A szuperédes fajták három-négyszeres mennyiségben tartalmaznak szacharózt a normáledes fajtákhoz képest. A glükóz, mint egyszerű cukor, a szervezet közvetlen energiaforrása, a leggyorsabban hasznosítható energiaadó vegyület, ennek megfelelően glikémiás indexe a legmagasabb az összes szénhidrát közül. A glükóz édessége hozzávetőlegesen háromnegyede a szacharózénak. A fruktóz a legtöbb gyümölcsben és zöldségben megtalálható. Lassú hasznosulása miatt lassabban emeli a vércukorszintet, glikémiás indexe a legalacsonyabb a cukrok közül, ezen kívül édesítő képessége 1,2-1,8-szorosa a szacharózénak, így azonos édesítőhatás eléréséhez fruktózból kisebb mennyiségre van szükség [15]. A gyorsfagyasztott csemegekukorica egész évben értékes tápanyagokat biztosít a fogyasztók számára. A csemegekukorica beltartalmi értékei átlagosan 100 gramm szemre vannak vonatkoztatva [16] (1. táblázat).

A csemegekukorica termesztéséhez Magyarország éghajlati agro-ökológiai viszonyai kedvező feltételeket biztosítanak, ugyanakkor a klímaváltozás hatásait nem lehet figyelmen kívül hagyni a biztonságos termesztéshez. A Magyarországon megtermelt csemegekukorica feldolgozási arányai évek óta közel állandónak tekinthető. A csemegekukorica 99%-a ipari feldolgozású, amelynek kétharmadát a konzervipar, egyharmadát a hűtőipar dolgozza fel gyorsfagyasztott termékként. A termés 1%-a friss fogyasztású. A csemegekukorica azok közé a hazai ipari növények közé tartozik, amelyek jelenleg világviszonylatban is versenyképesek. Az elmúlt évtizedben a gyorsfagyasztott csemegekukorica exportmennyiségét tekintve Magyarország az elsők között szerepelt, ugyanakkor fontos kiemelni, hogy a nemzetközi adatbázisokban a kínai termelési és kereskedelmi adatok nem, vagy csak becsült értékekkel szerepelnek [17].

A nemzetközi és hazai csemegekukorica-nemesítés jellemzően a termesztés és a feldolgozóipar szempontjait figyelembe véve hajította végre fajtaszelektív tevékenységét. Ugyanakkor a csemegekukorica-termékpálya résztvevőinek – termesztők, feldolgozók, kereskedők, fogyasztók – fajtákkal szembeni igényei jelentősen eltérőek. A csemegekukorica-termesztés fő célja a jövedelmezőség növelése, ennek megfelelően a termesztők fő szempontjai a következők: hektáronkénti termésátlag, a kártevők elleni rezisztencia, szakaszolhatóság, csőkihozatal, csőhossz, termőképesség, szárazságtűrés, szárszilárdság, szemsorszámm növelése, éghajlati változások kitettségére való alkalmazkodási képesség és terméshibafokozás, érésidő csökkentése. A feldolgozók legfőbb szempontja a hatékony feldolgozhatóság: zsengeség, egyöntetűség (kukoricacső-egyenesség, szemsor-egyenesség), szemkihozatali arány, morzsolhatóság, technológiai folyamatok során alkalmazott gépek hatékonyságának növelése. Jelenleg a feldolgozóiparban mindösszesen 10-15 ipari és ter-

1. táblázat. Csemegekukorica átlagos beltartalmi értékei 100 gramm friss szemre számítva [16]
Table 1 Average nutritional values of sweet corn per 100 grams of fresh kernels [16]

Tápelemek Nutrients	Vitaminok Vitamins	Aminosavak Amino acids	Ásványi anyagok Minerals	Nyomelemek Trace elements
víz / water	E vitamin (tokoferol) vitamin E (tocopherol)	izoleucin	K	Fe
szénhidrát carbohydrate	C-vitamin vitamin C	leucin	Ca	Zn
szacharóz sucrose	karotin (A provitamin) carotene (A provitamin)	lizin	Na	Cu
fruktóz fructose	tiamin thiamine	metionin	Mg	Mn
glükóz glucose	riboflavin riboflavin	fenil-alanin	P	
nyersrost raw fiber	niacin niacin	treonin		
fehérje protein	pantoténsav pantothenic acid	triptofán		
zsír fat	folsav folic acid	valin		
	B6-vitamin (Piridoxin) vitamin B6 (Pyridoxine)			



A kép illusztráció / Picture is for illustration only
Fotó/Photo: Shutterstock

mesztési tulajdonságokra optimalizált csemegekukorica-fajta a legnépszerűbb. Az előbbieken felsorolt tulajdonságok elsősorban genetikailag kódoltak a fajták génjeiben, azonban a környezeti körülmények – ökológiai viszonyok, agrotechnikai műveletek – befolyásolják ezeket a tulajdonságokat. A kereskedők legfőbb szempontja a profitmaximalizálás és a piacon érvényesíthető termékélnyök: egyöntetű szín, szemnagyság, íz- és állományjellemzők. Általánosan bevett gyakorlat a kereskedelembe, hogy a partnerek az elküldött minták alapján érzékszervi vizsgálat alapján döntenek [15] a felajánlott termények megvásárlásáról.

3. Célkitűzés

A hazai kutatások fókuszában ez idáig kifejezetten a feldolgozóipari minőség, a növényi és agronómiai tulajdonságok valamint a betegség-ellenállóság voltak [18]. A nemzeti és nemzetközi fajtajegyzékben található fajták érzékszervi oldalról történő komplex értékelése eddig nem valósult meg. Még kevesebb ismeretünk van a gyorsfagyasztott kukorica-hibridek, fajtákra lebontott érzékszervi profiljellemzőiről, fogyasztói preferenciáiról. A fogyasztói igények közepontba állításával történő, fajtákra lebontott részletes vizsgálatokat eddig nem publikáltak. Vizsgálataink során ezért célul tűztük ki az egyes fagyasztott csemegekukorica-fajták szakértői érzékszervi profiljait, valamint fogyasztói kedveltsége alapján felépített neurális hálózatok segítségével további fajták kedveltségének előrejelzésére nyíljon lehetőségünk.

4. Anyag és módszer

4.1. A vizsgálatba vont kukoricafajták

A kutatás tárgyát a csemegekukorica-fajtákból készült gyorsfagyasztott mintái jelentették. A kutatásainkba 41 csemegekukorica-fajta került, a minták a fajták megnevezésével jelöltük. A kiválasztott fajták egy része az évek során bizonyította folyamatos termesztésre való alkalmasságát, nagyobb részük viszont kisebb részaránytalálható meg a feldolgozóiparban, érzékszervi tulajdonságairól, kedveltségükről, kevés ismerettel rendelkezünk (2. táblázat).

A minták előkészítését minden esetben egyformán végeztük (főzési idő, edényzet nagysága, anyaga, márkája, főzőlap nagysága és hőmérséklete, vízmennyiség stb.). A minták szervóvizsgálásánál figyelembe vettük továbbá Kilcast [19] ajánlásait, miszerint egy személy készítette elő a kis mintamennyiségeket a jobb homogenitás érdekében. Minden bíráló azonos edényzetben 100 g azonos hőmérsékletű mintát értékelt. A mintákat a nemzetközi gyakorlatoknak megfelelően véletlen számgenerátorral előállított számhármassal kódoltuk [20]. A minták közötti ízsemlegesítéshez a szakirodalomban a termék jellegétől függően különböző élelmiszereket alkalmaznak, a tesztünkönk semleges jellegű ásványvizet használtunk [21].

4.2 Szakértői profilanalízis módszere

A profilanalízis módszere az egyik legösszetettebb érzékszervi vizsgálat, amely teljeskörűen (szín, íz, illat, állomány) jellemzi az adott élelmiszert. A bírálók a minták minősítését érzékszervi tulajdonságok mentén skálák segítségével teszik meg. A minősítéshez a bírálóbizottság tagjai két lépcsőben, először egyéni, majd közös munkával határozták meg a leíró kifejezések körét. A módszer esetében a bírálóbizottság tagjai jellemzően számos tulajdonságot értékelnek [22], [23].

A kísérlet megtervezését, bevonandó termékek és bírálók számának meghatározását, végrehajtását és az eredmények értékelését a vonatkozó szabványok előírásai alapján végeztük [24], [25]. Kutatásunkban a bírálók – a Budapesti Műszaki és Gazdaságtudományi Egyetem, Biokémia és Élelmiszertechnológia Tanszék, valamint a Szent István Egyetem, Érzékszervi Minősítő Laboratórium által fejlesztett – ProfiSens célszoftverrel értékelték a kísérleti mintákat. A szoftver segítségével néhány párbeszédablak kitöltésével elkészíthető a bírálati lap, a minták kiosztása (kitchen list) és a minták kiosztásához szükséges „tálca-alátét”. A minősítés ennek megfelelően a következőkben vázolt lépések szerint történt:

A bírálók 3 jegyű kódokkal ellátott mintasort kaptak, amelyet a közösen létrehozott rendszerben tulajdonságoként skálákon értékelték. A 16 vizsgálati szempontot a képzett bírálói panel határozta meg. A bírálók az értékelést 0-100-ig terjedő strukturálatlan skálán hajtották végre, amelyeknek két szélsőértékeit konszenzussal állapították meg. Az adott pontszámok szórásainak csökkentése érdekében lehorgonyozták a skálát és az egyik kukorica fajtának (‘Royalty’) tulajdonságoként megállapították a referenciaértékeit, amelyek a következők voltak: sárga szín (60), árnyalat (85), szemméret (55), szemméret egyenetlensége (80), frissesség (85), globális illatintenzitás (70), főtt-kukorica-illat (85), édes illat (70), állomány (75), lédús-ság (75), héj rághatósága (85), zsengesség (45), globális ízintenzitás (40), édes íz (35), főtt íz (20), utóíz (0).

A bírálókat helyi hálózatba szervezett számítógépekkel történtek, egymástól elszeparált bírálói fülkékben. A bírálókat után a ProfiSens szoftver olvassa ki a kitöltött elektronikus bírálati lapokból az egyes mintákra és tulajdonságokra vonatkozó bírálati eredményeket. Az eredmények statisztikai értékeléseként megkapjuk a fajtákhoz tartozó profildiógramot, illetve a tulajdonságoként átlagérték és szórások mellett egytényezős varianciaanalízis készült, ahol szignifikáns különbség adódott, ott két különböző valószínűségi szinten ($p=5\%$ és $p=1\%$) páronkénti *post hoc* tesztet végeztünk. A szakértői bírálókat egymást követő két napon, délelőtt 10 órakor kezdődtek, így két ismétlést hajtottunk végre. A szakértői panel tagjai „képzett bírálók” minősítési szinttel és tapasztalattal rendelkeztek (14 fő). A bírálókat végző személyek évek óta tagjai a Szent István Egyetem, Érzékszervi Minősítő

Laboratórium paneljének, nagy tapasztalattal rendelkeztek mind a módszert, mind a szoftver használatát illetően, és hasonló teszteken, termékspecifikus vizsgálatokban is rendszeresen részt vettek [26]. A vizsgálatok a nemzetközi irányelveknek megfelelően kialakított Szent István Egyetem, Érzékszervi Minősítő Laboratóriumában történtek [27].

4.3 Fogyasztói kedveltségvizsgálat

A fajták érzékszervi profiljellemzői alapján határoztuk meg, hogy a 41 fajta közül melyik 6 fajtát teszteljük a fogyasztók. Ezért az előzőekben elvégzett szakértői bírálókat adataira klaszterelemzést (Agglomerative Hierarchical Clustering, AHC) végeztünk Euklideszi távolság, Ward módszerrel. A klaszterezés során a szakértő bírálók eredményeit átlagoltuk, így előállítva a bemeneti terméktulajdonság \times fajta mátrixot. Az optimális klaszterszámot Silhouette-index alapján határoztuk meg, amely a legmagasabb értéket a hat klaszteres megoldásnál adta [28]. Az így kapott klaszterek közül a rangszámösszegek különbsége (sum of ranking differences, SRD) módszer segítségével klaszterenként meghatároztuk azokat a fajtákat, amelyek a legjobban leírják az adott klasztert [29], [30]. Így megkaptuk azt a hat fajtát (klaszterenként egyet), amely a „legátlagosabb” a klaszterekben. A fogyasztói panellel a kapott hat minta érzékszervi bírálókat végeztük el.

2. táblázat. A vizsgált kukoricafajták és fenntartói listája
Table 2 The investigated corn varieties and their maintainers

Sorszám Serial no.	Név (fenntartó) Name (maintainer)	Sorszám Serial no.	Név (fenntartó) Name (maintainer)	Sorszám Serial no.	Név (fenntartó) Name (maintainer)
1	‘Basin R’ (SVS Holland BV)	15	‘GSS 8529’ (Syngenta Seeds BV)	29	‘Rebecca’ (Pop Vriend BV)
2	‘Boston’ (Syngenta Seeds BV)	16	‘Jubilee’ (Syngenta Seeds BV)	30	‘Rocket’ (Harris Moran Seeds Co)
3	‘Box R’ (Topcorn Kft.)	17	‘Jumbo’ (Crookham Co.)	31	‘Royalty’ (Pop Vriend BV)
4	‘Dessert 82’ (Topcorn Kft.)	18	‘Jurassic’ (Syngenta Seeds BV)	32	‘Rustler’ (Vilmorin-NL)
5	‘Dessert R68’ (Topcorn Kft.)	19	‘Kinze’ (HM Clause SA)	33	‘SC 1036’ (Semini)
6	‘Dessert R78’ (Topcorn Kft.)	20	‘Kuatour’ (Harris Moran Seeds Co.)	34	‘Sheba’ (Asgrow Seeds Co)
7	‘Dynamo’ (Harris Moran Seeds Co.)	21	‘Legend’ (HM Clause SA)	35	‘Spirit’ (Syngenta Seeds BV)
8	‘Enterprise’ (Snowy River Seeds Ply Ltd.)	22	‘Madonna’ (SVS Holland BV)	36	‘Shinrock’ (Syngenta Seeds BV)
9	‘Galaxy’ (HM Clause SA)	23	‘Mercur’ (Royal Sluis BV)	37	‘Starshine’ (Syngenta Seeds BV)
10	‘Garrison’ (Syngenta Seeds BV)	24	‘Merit’ (Royal Sluis BV)	38	‘Sweetstar’ (Syngenta Seeds BV)
11	‘GH 2042’ (Syngenta Seeds BV)	25	‘Noa’ (Pop Vriend BV)	39	‘Tasty Sweet’ (IFS Inc.)
12	‘GH 6225’ (Syngenta Seeds BV)	26	‘Overland’ (Syngenta Seeds BV)	40	‘TOP 825’ (Topcorn Kft.)
13	‘GSS 1477’ (Syngenta Seeds BV)	27	‘Prelude’ (Snowy River Seeds Ply Ltd.)	41	‘Turbo’ (Harris Moran Seeds Co.)
14	‘GSS 5649’ (Syngenta Seeds BV)	28	‘Puma’ (Crookham Co.)		

A laikus fogyasztói bírálók (167 fő) ezt a 6 mintát tesztelték, akik kizárólag a skálák és a szoftver használatával kapcsolatban kaptak információt. A termékhez kötődően semmilyen speciális képzettséggel nem rendelkeztek sem gyakorlati, sem elméleti szempontból, továbbá érzékszerveik érzékenységét sem vizsgáltuk. A fogyasztók a termékek globális kedveltségére adtak választ egy 9 elemű strukturált, folyamatosan növekvő skálán (1=egyáltalán nem, 2=nagyon nem, 3=mérsékelten nem, 4=kissé nem, 5=közömbös, 6=kissé kedvelt, 7=mérsékelten, 8=nagyon kedvelt, 9=mindennél jobban).

4.4 Alkalmazott mesterséges neurális hálózatok

A kutatásainkat a Palisade szoftvercsalád Neural Tools ver. 5.5 szoftverrel végeztük. A partícionálás során egy 1000-szeres Monte Carlo szimuláció adatainak 80 %-án tréningeztük a modellt, majd a maradék 20 % szolgáltatta a tesztfuttatások kiinduló adatait. Az többrétegű előreccsatolt (multi-layer feedforward neural net, MLFN) háló struktúrájának optimalizálásához a „Best Net Search” lehetőséget választottuk, amely öt hálót tesztel 2-6 nódusszal, és kiválasztja a legjobb predikciót adott. A NeuralTools „Best Net Search” opcióját a túltréningezés megelőzésére alakították ki. Alapbeállításokkal a „Best Net Search” 2 neuronnal kezd egy hálót tesztelni, ami tipikusan túl kicsi, hogy túltréningezzük. Alapbeállításokkal egé-

szen 6 neuronig fog hálókat tréningezni. Ha az 5 és 6 neuronos hálókat túltréningezi, akkor az megjelenik az eredményekben. A 2, 3 vagy 4 neuronos hálók egyikének lesz a legalacsonyabb teszthibája.

5. Eredmények

A „Best Net Search” beállítás segítségével a szoftver hat MLFN konfigurációt tesztelt, amíg kiválasztotta a legjobb predikciót adó konfigurációt (**1. ábra**). A fogyasztók esetében a 4 nódusos MLFN adta a legjobb eredményeket (**3. táblázat**).

Az MLFN-modell kialakításánál a tréningezéshez a mintákat véletlenszerűen választottuk ki. A modellek maradékai megadják, hogy a szakértői adatokból

háló a fogyasztói kedveltséget milyen pontossággal tudta előre jelezni. A legkisebb maradékokat a tréning és a teszt során is a 4 nódusosból álló háló adta. A kapott neurális háló validitását a maradékok értelmezése mellett véletlen számokon történő előrejelzéssel is vizsgáltuk. Az így kapott eredmények alapján bebizonyosodott, hogy a háló a véletlen számokat rosszul jelezte előre, nem talált összefüggést az adatok között. A fentiek mellett a modell keresztellenőrzését is elvégeztük, amely alapján nem mutatott szignifikáns eltéréseket. Az ellenőrzés során a háló 87 %-os pontossági értéket adott. Ezek alapján a modellt elfogadtuk, és a továbbiakban a 4 nódusos MLP modellt használtuk a többi fajta szakértői adatokból történő kedveltségének előrejelzésére.

3. táblázat. A hálózatok teszteléseinek „Best Net Search” eredményei
Table 3 Results of „Best Net Search” network tests

Best Net Search Best Net Search	Minimum Maradék (tréning) Minimum residue (training)	Maximum Maradék (tréning) Maximum residue (training)	Minimum Maradék (teszt) Minimum residue (test)	Maximum Maradék (teszt) Maximum residue (test)
MLFN 2 nódus MLFN 2-node	-1.06007	0.96618	-0.87118	0.99700
MLFN 3 nódus MLFN 3-node	-0.97762	1.03442	-0.88366	0.99625
MLFN 4 nódus MLFN 4-node	-0.95308	0.96563	-0.88233	0.98695
MLFN 5 nódus MLFN 5-node	-1.03178	1.0418	-0.98502	1.10380
MLFN 6 nódus MLFN 6-node	-1.11483	1.15705	-1.05110	1.14759

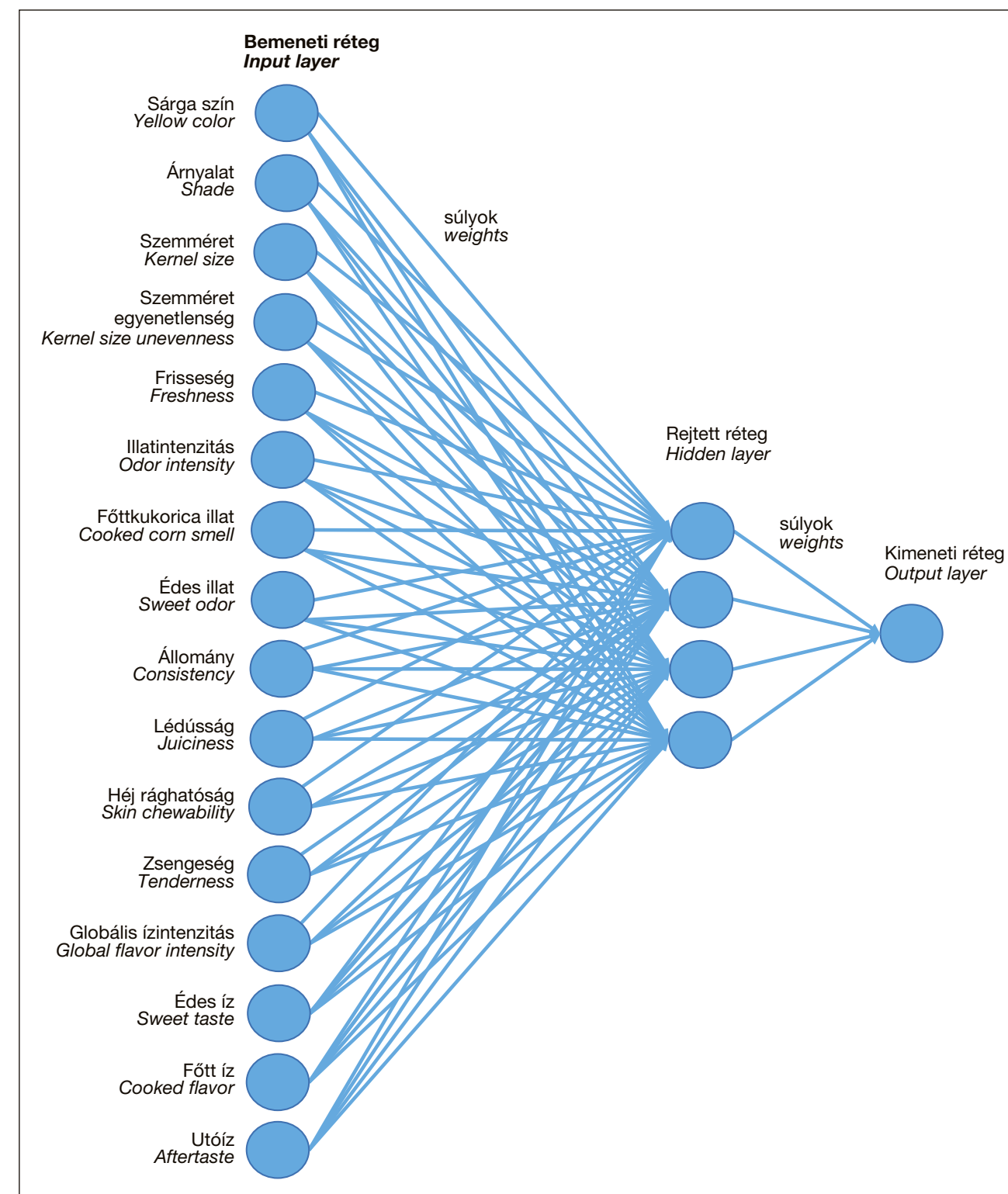
4. táblázat. A 4 nódusos MLP háló előrejelzései a fogyasztói kedveltségre a 9-tagú kategóriaskálán
Table 4 Consumer preference predictions of the 4-node MLP net on a 9-point category scale

Fajta neve Variety	Prediktált kedveltségi érték Predicted preference value	Fajta neve Variety	Prediktált kedveltségi érték Predicted preference value
'Basin R'	6.80	'Merit'	2.80
'Boston'	7.63	'Noa'	6.83
'Dessert 82'	6.80	'Overland'	7.16
'Dessert R78'	6.80	'Prelude'	4.46
'Enterprise'	4.46	'Puma'	4.21
'Garrison'	8.10	'Rebecca'	7.14
'GH 6225'	4.46	'Rocket'	3.16
'SC 1036'	6.81	' Rustler '	2.70
'GH 2042'	3.16	'Sheba'	8.00
'GSS 8529'	7.16	' Shinerock '	8.46
'GSS 1477'	8.04	'Starshine'	3.78
'GSS 5649'	7.64	'Sweetstar'	8.09
'Jubilee'	4.46	'Tasty Sweet'	6.80
'Jumbo'	2.69	'TOP 825'	6.39
'Kinze'	8.10	'Box R'	3.63
'Kuatour'	2.79	'Dessert R68'	7.16
'Legend'	3.16	'Turbo'	5.44
'Madonna'	2.70		

A felépített modell a 9 tagú skálán a leginkább kedveltnek a 'Shinerock' fajtát (8,46) adja meg, míg a predikciók szerint 2,7-es átlagos kedveltségi értékkel a legkevésbé kedvelt a 'Madonna' és a 'Rustler' fajták lettek (**4. táblázat**).

A Palisade szoftver a neurális hálózatok tréningezése és tesztelése során a háló felépítésében szerepet játszó változók fontossági sorrendjét is megadja (**2. ábra**).

A csemegekukorica-fajták esetében a szakértői adatok és a fogyasztói kedveltség-adatok közti összefüggés keresése közben a legfontosabb változóknak az édes íz (18 %), a globális ízintenzitás (14 %) és a lédúság (12 %) adódtak. A hálózat eredményei alapján tehát a fogyasztók az intenzív édes ízű és lédús termékek értékelésekor adtak magasabb kedveltségi pontszámokat. Az eredmények alátámasztják a korábbi kutatások eredményeit [31], [32], [33].



1. ábra. Mesterséges neurális háló (4 nódusos MLFN) összefüggésrendszere
Figure 1 Correlation system of an artificial neural network (4-node MLFN)

Az így kapott eredményeket ezt követően összehasonlítottuk azokkal a kezdeti klaszterekkel, amelyekből a „legátlagosabb” (a klaszter tagjait legjobban reprezentáló) fajta kiválasztása után a neurális hálózatot felépítettük. Az **5. táblázat** eredményei alapján látható, hogy az első két klaszter édesíz-intenzitása szignifikánsan nagyobb értékekkel rendelkezik, mint a többi klaszter (az 5. klasztertől nem egyértelmű az elkülönülés).

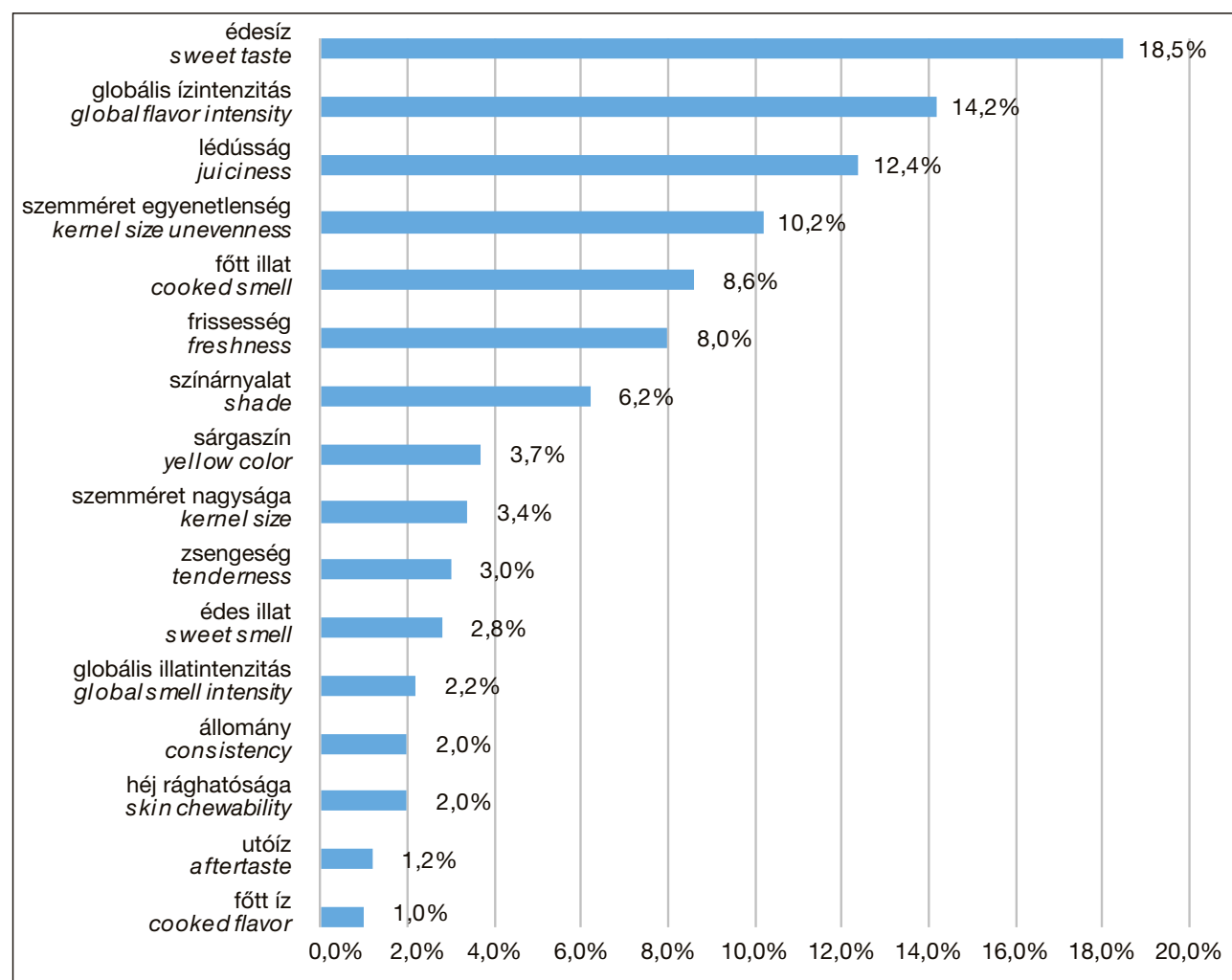
Hasonlóan nagyobb értékeket kapott az első két klaszter a globális ízintenzitás értékeire is, azonban itt a 3. és 6. klaszter is magas értékekkel szerepel, így a 4. és 5. klasztertől ezek szignifikánsan különböztek a Tukey-HSD próba eredményei alapján. Fontos megemlíteni, hogy a globális ízintenzitás értékelése során a bírálók az összes íz intenzitása alapján értékelnek, amely ízek nem feltétlenül jelentenek a fogyasztói értékelés során előnyöket. Zsengeség alapján az első két klaszter mellett a 6. klaszter tagjai szignifikánsan magasabb értékekkel rendelkeznek, mint a többi klaszter, így két csoportra bontottuk a vizsgált klasztereket (**6. táblázat**).

Az eredmények alapján látható, hogy az első két klaszter intenzív globális és édes ízekkel rendelkezik, illetve határozottan zsenge termékek voltak. Az ered-

ményeket a prediktált fogyasztói eredmények is tükrözik, mivel az első két klaszter szinte azonos eredményeket, 6,8-as kedveltségi átlagértékeket kapott a neurális hálózat alapján. A harmadik klaszter átlagos kedveltségi értéke 5,2-nek adódott, amely nagymértékben köszönhető a magas (80) globális ízintenzitás értéknek. A következő csoportot a 4,0 és 4,3-as kedveltségi átlagértékekkel rendelkező negyedik és hatodik klaszterek alkotják. A hatodik klaszter tagjai magasabb zsengeség- és globális ízintenzitás-értékekkel rendelkeznek, míg a negyedik klaszter tagjainak zsengeség-értéke közepes eredményt jelent. A legkevésbé kedvelt mintákat az ötödik klaszter tartalmazza, amelyek alacsony intenzitásértékekkel rendelkeznek mindhárom, a kedveltség előrejelzésénél fontos terméktulajdonságnál.

6. Következtetések

A bemutatott mesterséges neurális hálózatokat és Monte Carlo szimulációt kombináló megközelítés alkalmazásnak bizonyult a szakértői érzékszervi bírálati eredmények alapján a fogyasztói kedveltség előrejelzésére. A megközelítés előnye, hogy a sok időt, energiát, költséget felemésztő fogyasztói vizsgálatok eredményesen kiválthatóak a szakértői adatok alapján végzett predikcióval. Összesen 36 csemegekuko-



2. ábra. A neurális háló felépítésében résztvevő változók relatív fontossága és sorrendje
Figure 2 Relative importance and order of variables playing a role in the structure of the neural network

rica-minta kedveltségi értékeit sikerrel jeleztük előre hat minta fogyasztói értékelésének eredményei alapján. A mesterséges neurális hálózatmodell megalkotása során sikeresen azonosítottuk azokat a terméktulajdonságokat, amelyek a fogyasztói elfogadás fő mozgatórugói. Ezek fontossági sorrendben: az édes íz, a globális ízintenzitás és a lédúság voltak. Kutatásunkat a jövőben célszerű lenne további kertészeti és élelmiszeripari termékekre is kiterjeszteni, illetve olyan termékek esetében is felhasználni, amelyeknél a fogyasztói preferenciát nem néhány, jól behatárol-

ható termékjellemző befolyásolja. További lehetőség egy olyan szoftveres megvalósítás létrehozása, amely leegyszerűsíti a számítási lépéseket, így ezáltal egy szoftveren belül elvégezhető lenne a hálózatok tesztelése, validációja és az új adatok predikciója.

Összefoglalóan megállapítható, hogy a validált termékspecifikus mesterséges neurális hálózatok lehetővé tehetik a legfontosabb érzékszervi tulajdonságok meghatározását. Az új megközelítés hatására az eredmények megbízhatóbbak, az ismétlések könny-

5. táblázat. A három legfontosabb szakértői érzékszervi terméktulajdonság klaszterenkénti átlagértékei összevetve a neurális háló modell által prediktált klaszterenkénti kedveltségértékek átlagával és szórásával

Table 5 Average values of the three most important expert sensory product properties for each cluster compared to the average values and standard deviations of the preference values predicted for the clusters by the neural network model

	édes íz Sweet taste	globális ízintenzitás Global flavor intensity	zsengeség Tenderness	kedveltség átlag Preference average	kedveltség szórás Preference standard deviation
1. klaszter Cluster 1	63.05	63.35	69.28	6.83	1.90
2. klaszter Cluster 2	75.54	73.05	61.97	6.84	1.21
3. klaszter Cluster 3	23.79	80.97	13.07	5.26	2.41
4. klaszter Cluster 4	24.55	27.92	35.77	4.04	0.67
5. klaszter Cluster 5	26.72	21.77	15.03	3.28	0.65
6. klaszter Cluster 6	21.22	62.50	65.20	4.33	2.12

6. táblázat. Az egyes terméktulajdonságok klaszterenkénti összehasonlítása, homogén és heterogén csoportok (Tukey-HSD próba)

Table 6 Comparison of different product properties by cluster, homogeneous and heterogeneous groups (Tukey-HSD test)

Klaszterek Cluster	Zsengeség-átlag Tenderness average	Csoportok Groups	
1	69.28	A	
6	65.20	A	
2	61.97	A	
4	35.77		B
5	15.03		B
3	13.07		B
Klaszterek Cluster	Globális ízintenzitás-átlag Global flavor intensity average	Csoportok Groups	
3	80.97	A	
2	73.05	A	B
1	63.35	A	B
6	62.50		B
4	27.92		C
5	21.77		C
Klaszterek Cluster	Édes íz intenzitás-átlag Sweet taste intensity average	Csoportok Groups	
2	75.54	A	
1	63.06	A	B
5	26.72		B C
4	24.55		C
3	23.79		C
6	21.22		C

nyebben végrehajthatók, a vizsgálatok jobban reprodukálhatók, összességében egy idő- és költségkímélő elemzési rendszer jön létre.

7. Köszönetnyilvánítás



Kutatásainkat, munkánkat az Emberi Erőforrások Minisztériuma ÚNKP-17-4 kódszámú Új Nemzeti Kiválóság Programja, a Bolyai János kutatási ösztöndíj, valamint az OTKA K112547 számú pályázat támogatásával végeztük. A támogatásokat köszönjük.

8. Irodalom

- [1] McCulloch, W. S. and Pitts, W. H. (1943): A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115–137.
- [2] Hebb, D.O. (1949): *The organization of behavior: A neuropsychological theory*. New York, John Wiley and Sons. 335.
- [3] Rosenblatt, F. (1958): The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65 (6), 386–408.



A kép illusztráció / Picture is for illustration only
Fotó/Photo: Shutterstock

- [4] Widrow, B. and Hoff, M.E. Jr. (1960): Adaptive Switching Circuits. IRE WESCON Convention Record. 96–104.
- [5] Borosy, A. P. (2001): Mesterséges ideghálózatok. In: Horváth szerk. *Sokváltozós adatelemzés (Kemometria)*. Budapest, Nemzeti tankönyvkiadó. 312–329.
- [6] Hopfield, J. (1982): Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci. U.S.A.*, 79 (8) 2554–2558.
- [7] Rumelhart, D.E. and McClelland, J.L., eds (1986) *Parallel Distributed Processing, Explorations in the Microstructure of Cognition*, MIT Press.
- [8] Horváth G. (szerk.) (2006): *Neurális hálózatok*. Altrichter M., Horváth G., Pataki B., Strausz Gy., Takács G., Vályon J., Neurális hálózatok, Budapest, Panem Kiadó, 2006.
- [9] Sipos László, Gere Attila, Kókai Zoltán, Szabó Dániel (2012): Mesterséges ideghálózatok (ANN) alkalmazása az érzékszervi minősítés gyakorlatában. 58, (1-2), 32-46.
- [10] Marini, F. (2009): Artificial neural networks in foodstuff analyses: Trends and perspectives. *A review. Analytica Chimica Acta*, 635, 121–131.
- [11] Debska B., Guzowska-Swider, B. (2011): Application of artificial neural network in food classification. *Analytica Chimica Acta*, 705, 283–291.
- [12] *NeuralTools Version 5.7 Manual* (2010).
- [13] Fu, LiMin (1994): *Neural Networks in Computer Intelligence*. McGraw-Hill, Inc. pp. 1-416.
- [14] Székely Géza, Losó Viktor, Tóth Arnold: *Nemzetközi és hazai zöldség-gyümölcsfogyasztás, módszertani kérdések- ÉVIK 2015/1*
- [15] Losó Viktor (2015): Gyorsfagyasztott csemegekukorica termékek komplex értékelése. *Doktori Értekezés*. Budapesti Corvinus Egyetem. 7-27.
- [16] Bíró, Gy., Lindner K. (Szerk.) (1999): *Tápanyagtáblázat: Táplálkozásban és tápanyagösszetétel*, Budapest, Medicina Kiadó.
- [17] Fodor, Z. (szerk.) (2016): *Zöldség és gyümölcs ágazat helyzete Magyarországon*. Budapest, FruitVeB Magyar Zöldség-Gyümölcs Szakmaközi Szervezet. 1-31.
- [18] Orosz, F. (2009): *Termesztéstechnológiai elemek hatása acsemegekukorica koraiságára*. *Doktori értekezés*, Budapest. p. 1- 157.
- [19] Kilcast, D. (2010): *Sensory analysis for food and beverage quality control*. Woodhead, Cambridge, UK.
- [20] MSZ ISO 6658:2007 *Érzékszervi vizsgálat. Módszertan. Általános útmutató*
- [21] Sipos, L. (2009): *Ásványvízfogyasztási szokások elemzése és ásványvizek érzékszervi vizsgálata*. *Doktori értekezés*. Budapesti Corvinus Egyetem. 80-102.
- [22] Kókai, Z. (2003): *Az almafajták érzékszervi bírálata*. *Doktori értekezés*. Budapest: Budapesti Közgazdaságtudományi és Államigazgatási Egyetem, 35-59.
- [23] Varela, P., Ares, G. (2014): *Novel Techniques in Sensory Characterization and Consumer Profiling*, CRC Press, pp. 9-41.
- [24] MSZ ISO 11035:2001 *Érzékszervi vizsgálat. A leíró kifejezések azonosítása és kiválasztása érzékszervi profilhoz többdimenziós eljárással*
- [25] ISO 13299:2003 *Sensory analysis – Methodology – General guidance for establishing a sensory profile*
- [26] MSZ EN ISO 8586:2014 *Érzékszervi vizsgálat. Általános útmutató a kiválasztott bírálók és az érzékszervi szakértő bírálók kiválasztásához, képzéséhez, valamint folyamatos ellenőrzéséhez*
- [27] MSZ EN ISO 8589:2015 *Érzékszervi vizsgálatok. Általános útmutató a bírálói helyiségek kialakításához*
- [28] Chen G.X, Jaradat S.A, Banerjee N., Tanaka T.S., Ko M.S.H., Zhang M.Q. (2002): Evaluation and comparison of clustering algorithms in analyzing ES cell gene expression data. *Stat Sin*, 12:241–262.
- [29] Héberger, K. (2010): Sum of ranking differences compares methods or models fairly. *Trend. Anal. Chem.* 29, 101–109.
- [30] Héberger, K., Kollár-Hunek, K. (2011): Sum of ranking differences for method discrimination and its validation: comparison of ranks with random numbers. *Journal of Chemometrics*, 25, (4) 151–158.
- [31] Gere, A., Losó, V., Tóth, A., Kókai, Z., Sipos, L. (2012): *Kukorica fajták preferenciaterképezése szoftveres támogatással. Élelmiszervizsgálati Közlemények*, 58, pp. 118-130.
- [32] Gere, A., Losó, V., Radványi, D., Juhász, R., Kókai, Z., Sipos, L. (2013): *Csemegekukorica fajták komplex értékelése. Élelmiszervizsgálati Közlemények*, 59, pp. 120-134.
- [33] Gere, A., Losó, V., Györey, A., Kovács, S., Huzsvai, L., Nábrádi, A. Kókai, Z., Sipos L. (2014): Applying parallel factor analysis and Tucker-3 methods on sensory and instrumental data to establish preference maps. Case study on sweet corn varieties. *Journal of the Science of Food and Agriculture*, 94, 15, pp. 3213-3225.

¹László Sipos, ¹Viktor Losó, ¹Ákos Nyitrai, ¹Zoltán Kókai, ¹Attila Gere

Received: 2017. July – Accepted: 2017. October

Prediction of sensory preference by artificial neural networks, using sweet corn varieties as an example

Keywords: prediction, consumer preference, multi-layer feedforward neural network (MLFN), sweet corn, nutritional parameters of sweet corn, cluster analysis

1. Summary

According to our knowledge, there are only a few publications in available literature sources on the sensory characteristics and consumer preferences of sweet corn varieties. In our research, practical application of artificial neural networks (ANNs) is presented. In our study, 41 frozen sweet corn varieties were evaluated by a panel of expert sensory panelists (14 persons), by the method of profile analysis (MSZ ISO 11035:2001; ISO 13299:2003), on an unstructured scale of 0 to 100, then, in large-scale tests, 6 of the 41 varieties were evaluated by consumers (167 people) according to preference, on a structured scale of 1 to 9.

Artificial neural networks require large amounts of data, therefore, on the expert and consumer data for the 6 varieties, 1,000 Monte Carlo simulations were run. 80% of the resulting dataset was used to train the created neural networks, and 20% was utilized to test them. The best prediction was given by the 4-node multi-layer feedforward neural network (MLFN), the smallest residues were obtained in this case during the training and the test, which were also validated by predictions on random numbers and cross-checking. Preference values of the other 35 corn varieties were predicted by this model. The most preferred variety was ‘Shinerock’ (8.46), while the least preferred ones, according to the predictions, were ‘Madonna’ and ‘Rustler’, with an average preference value of 2.7 (on a scale of 1 to 9).

During the establishment of the artificial neural network model, product characteristics that are the main drivers of consumer acceptance were successfully identified: sweet taste, global taste intensity and juiciness. In general, it can be stated that prediction of the preference of different varieties is made possible by the validated product-specific artificial neural network presented.

2. Introduction and literature overview

For the development of artificial neural networks, revealing the analogy, the structure and functioning of the human nervous system was of key importance. Neural network programs were originally developed as a model for the nervous system, where signals coming from other neurons are collected by the in-

puts, summation is carried out by the processing unit (neuron), and then, depending on the result, the signal is transmitted by the outputs [1], [2], [3], [4], [5]. A breakthrough in the research of artificial neural networks was achieved by the work of Hopfield [6], and Rumelhart et al. [7], in which non-linear mapping was achieved by the dynamic modeling of neural network programs, as well as feedback between the outputs and the inputs. Artificial neural networks in today’s

sense are network systems organized by the connection of simple processing units (neurons) capable of parallel operation, possessing learning and information recall algorithms.

Artificial neural networks are systems that learn from past experience, and are able to recall the things learned. Their main advantages are non-linearity, good fit, parallel calculations, high calculation speed and adaptability. The most important application areas are the identification of outliers, identification of the correlation between variables, space reduction, regression between nonlinear variables, modeling of complex relationships, classification and categorization [8], [9].

The most important parts of neural networks are the elements making up the network (neurons/nodes), network connections (structure/topology) and learning algorithms. The operating principle of ANN is based on the modeling of functional connection between the input and corresponding output variables: $y = f(x)$, where x and y are the input and output vectors, and the symbol f indicates that there is a functional connection. The signals received by the neurons are multiplied by the corresponding weighting factors, and then these are summed up. Next, the value of the output signal is calculated by the neuron, typically by a nonlinear transformation function – step, hyperbolic tangent, logistic, sigmoid, etc. These neuron units may be connected to the neurons of other layers, passing the information from one neuron to the next until the final output, where the input signal becomes an output signal [10].

Determining the number of layers and neurons is particularly important, because it determines whether the network will be able to learn the relationships between independent and dependent variables. The number of hidden layers and the number of neurons in the hidden layers depends on the complexity of the classification task and the amount of data. In general, networks that contain one hidden layer and two hidden neurons cannot be trained to a satisfactory level of error. Any linear or nonlinear problem can be solved with the help of a hidden layer and a sigmoid activation function. The use of two or more hidden layers will unnecessarily increase training time. Usually, a few neurons are enough in a hidden layer. The number of branches (neurons) in the input layer is in accordance with the number of variables describing the subjects classified, while the number of branches in the output layer increases with the number of classes [11], [12], [13].

Based on the learning methods of networks, we can distinguish between *supervised* and *unsupervised* networks. Based on the connections and structural system (topology) of supervised learning networks, they can be either *feedforward networks* or *recurrent networks*. In multi-layer feedforward networks, information flows forward in a way that neurons in

the given layer do not receive a signal until it is produced by the units in the previous layer, so the output vector of each layer is also the input vector of the next layer. Unsupervised learning networks include Kohonen-based structures [10]. Practical application of neural networks consists of three main steps. The neural network is built up during the learning phase, the validation, with model indicators, of the neural network created is performed in the validation step, and the applicability of the neural network can be investigated by the testing step [5], [11].

In the following paragraphs, the most relevant nutritional, agrotechnical and breeding issues of sweet corn will be presented. Vegetables greatly contribute to the supply of vitamins and minerals for the human body, they affect carbohydrate and fat metabolism, and all other processes in the body that are related to their fiber content [14]. Compared to the average vegetable, sweet corn has a high energy content. Its nutritional significance is primarily due to its carbohydrate and protein content. It is characteristic of the carbohydrate profile of sweet corn varieties that the relative weight of the disaccharide sucrose, made up from a glucose and a fructose unit, is generally the highest (85%), followed by glucose (10%) and fructose (5%). Super-sweet varieties contain three to four times more sucrose, compared to normal sweet varieties. Glucose, as a simple sugar, is the direct energy source of the body, the energy-producing compound that can be utilized the fastest and due to this, its glycemic index is the highest among all carbohydrates. The sweetness of glucose is roughly three quarters of that of sucrose. Fructose can be found in most fruits and vegetables. Due to its slow utilization, it increases blood sugar levels more slowly, its glycemic index is the lowest among sugars and, besides, its sweetening ability is 1.2 to 1.8 times higher than that of sucrose, so to achieve the same sweetening effect, less fructose is needed [15]. Fast-frozen sweet corn provides valuable nutrients for consumers throughout the year. Nutritional values of sweet corn are generally calculated for 100 grams of kernels [16] (Table 1).

Agro-ecological conditions in Hungary are favorable for the growing of sweet corn, however, the effects of climate change cannot be ignored. The rates of processing of sweet corn produced in Hungary have been almost constant for years. 99% of sweet corn is processed industrially, two thirds of which is processed by the canning industry, and one third by the refrigeration industry as quick-frozen products. 1% of the harvest is consumed fresh. Sweet corn is one of the domestic industrial plants that are competitive on a global level. Over the past decade, Hungary has been among the first countries of the volume of quick-frozen sweet corn exported, however, it is important to stress that Chinese production and trade data are not included in international databases, or only as estimates [17].

¹ Szent István University, Faculty of Food Science, Department of Postharvest and Sensory Evaluation

Varietal selection activities have typically been performed by international and domestic breeders with the considerations of cultivation and the processing industry in mind. At the same time, demands of the stakeholders of the sweet corn sector (growers, processors, merchants, consumers) are significantly different. The main goal of the cultivation of sweet corn is to increase profitability and, accordingly, the main considerations of growers are average yield per hectare, resistance to pests, planting period, ear yield, ear length, productivity, drought tolerance, stem strength, increase in kernel row number, ability to adapt to climate change, increasing crop safety, reducing ripening time. The most important consideration of processors is efficient processability: tenderness, homogeneity (ear straightness, kernel row straightness), kernel yield, shellability, increasing the efficiency of the machines used in technological processes. Currently, there are 10 to 15 varieties of sweet corn, optimized for industrial and cultivation properties, that are most popular in the processing industry. The above-mentioned properties are primarily genetically encoded in the genes of the varieties, however, these properties can be influenced by environmental conditions (ecological conditions, agrotechnical operations). The most important considerations of merchants are profit maximization and marketable product benefits: uniform color, kernel size, flavor and firmness characteristics. It is common practice in commerce that partners decide whether to buy the crops offered on the basis of sensory tests performed on samples sent by the supplier [15].

3. Objective

Domestic research usually focused on processing industrial quality, plant and agronomic properties, as well as disease resistance [18]. No complex evaluation of the sensory properties of varieties in the national and international catalogs of varieties has been carried out. Even less is known about the sensory characteristics and consumer preferences of quick-frozen corn hybrids. No detailed investigations, focusing on consumer demand and broken down by variety, have been published so far. Therefore, our investigations were aiming to predict preference values, with the help of neural networks created on the basis of the expert sensory profiles of certain frozen sweet corn varieties and their consumer preference, the preference for other varieties.

4. Materials and methods

4.1. Sweet corn varieties

The subjects of the research were quick-frozen samples of sweet corn varieties. 41 sweet corn varieties were included in the research, and samples were designated by the names of the varieties. Over the years, some of the selected varieties have demonstrated their ability to grow continuously, however, most of them have a smaller share in the processing

industry, and little is known about their sensory properties and popularity (Table 2).

Sample preparation was carried out in identical manner in each case (cooking time, container size, material, brand, hotplate size and temperature, water volume etc.). When serving the samples, the recommendations of Kilcast [19] were also taken into account, according to which, for better homogeneity, the serving samples were prepared by the same person. 100 g of the sample of the same temperature in identical containers were evaluated by each judge. In accordance with international practice, samples were encoded by three-digit numbers generated by a random number generator [20]. In the literature, different foods are used for flavor neutralization between the samples, depending on the nature of the product, in our test we used mineral water [21].

4.2 Expert profile analysis method

The method of profile analysis is one of the most complex sensory tests, entirely (color, taste, aroma, consistency) characterizing the given food. Samples are evaluated by the panelists along sensory properties, with the help of scales. For classification, descriptive terms were defined by members of the sensory panel in two stages, first individually, and then through group work. In the case of this method, typically several properties are evaluated by the members of the sensory panel [22], [23].

Planning and execution of the experiment, determination of the number of products to be included and of the panelists, and the evaluation of the results were carried out in accordance with the relevant standards [24], [25]. In our research, experimental samples were evaluated by the panelists using the ProfiSens sensory software developed by the Budapest University of Technology and Economics, Department of Biochemistry and Food Technology and Szent István University, Laboratory for Sensory Analysis. With the help of the software, by filling in a few dialog boxes, the evaluation sheet, the *kitchen list* and the sample codes for the distribution of the samples can be prepared. Accordingly, qualification was carried out through the following steps:

Panelists received a series of samples labeled with three-digit codes, and these were evaluated each attribute on scales. The 16 sensory attributes had been determined by the panel of expert panelists. Evaluation by the panelists was carried out on an unstructured scale of 0 to 100, the two extreme values of which were determined by consensus. To reduce the standard deviation of the given scores the scale was anchored, and the reference values of one of the corn varieties (*Royalty*) were determined for each property, which were as follows: yellow color (60), shade (85), kernel size (55), unevenness of kernel size (80), freshness (85), global smell intensity (70), cooked

corn smell (85), sweet smell (70), consistency (75), juiciness (75), skin chewability (85), tenderness (45), global flavor intensity (40), sweet taste (35), cooked flavor (20), aftertaste (0).

Evaluation were carried out on computers in a local area network, in sensory booths separated from each other. After the evaluation, results for the individual samples and properties were read from the filled-in electronic evaluation sheets by the ProfiSens software. After statistical evaluation of the results, the profile diagrams of the varieties were obtained and, in addition to the average value and the standard deviation for each property, one-way analysis of variance was carried out, and in cases where there was a significant difference, *post hoc* tests were performed at two different probability levels ($p=5\%$ and $p=1\%$) for each pair. Expert evaluations begin at 10 am on two consecutive days, so two repetitions were carried out. Members of the expert panel possessed „trained assessors” qualifications and experience (14 people). The people performing the evaluation had been members of the panel of the Laboratory for Sensory Analysis of Szent István University for several years, with extensive experience in the use of both the method and the software, and had been participating in similar tests and product-specific studies regularly [26]. Tests were carried out in the Laboratory for Sensory Analysis of Szent István University, established in accordance with international guidelines [27].

4.3 Consumer preference testing

The 6 of 41 varieties to be tested by consumers were selected on the basis of the sensory profile characteristics of the varieties. For this reason, cluster analysis (*Agglomerative Hierarchical Clustering*, AHC) was carried out for the expert evaluation data performed earlier, using Euclidean distance and the Ward method. During clustering, results of the expert panelists were averaged, thus creating the input product property \times variety matrix. The optimal cluster number was determined by the Silhouette index, which gave the highest value in the case of the six-cluster solution [28]. For each cluster thus obtained, the varieties that describe best the given cluster were determined by the *sum of ranking differences* (SRD) method [29], [30]. This way, the six “most average” varieties in the clusters (one for each cluster) were obtained. Sensory evaluation of the six samples obtained were carried out by the consumer panel.

The 6 samples were tested by by a consumer panel (167 people), who had received information only regarding the use of the scales and the software. They did not have any special qualifications, either practical or theoretical, related to the product, and the sensitivity of their senses had not been investigated either. Consumers provided answers regarding the global preference of the products on a 9-point, structured, continuously increasing scale (1=not at all,

2=very much not, 3=moderately not, 4=slightly not, 5=neutral, 6=slightly preferred, 7=moderately preferred, 8=very much preferred, 9=most preferred).

4.4 Artificial neural networks used

Our research was carried out using the Neural Tools ver. 5.5 software of the Palisade software family. During partitioning, the model was trained on 80% of the data of 1,000 Monte Carlo simulations, and starting data for the test runs were provided by the remaining 20%. For the optimization of the *multi-layer feedforward neural network* (MLFN) structure, the „Best Net Search” option was selected, which tests five nets with 2 to 6 nodes, and then selects the one giving the best prediction. The „Best Net Search” option of NeuralTools was developed to prevent overtraining. By default, „Best Net Search” starts to test a net with 2 neurons, which is typically too small to overtrain. With default settings, it will train nets with up to 6 neurons. If nets with 5 or 6 neurons are overtrained, it will appear in the results. The testing error of one of the nets with 2, 3 or 4 neurons will be the lowest.

5. Results

Using the „Best Net Search” setting, six MLFN configurations were tested by the software, until the configuration giving the best prediction was selected (Figure 1). In the case of consumers, best results were given by the 4-node MLFN (Table 3).

When developing the MLFN model, samples for training were selected at random. Residues of the model indicate how accurately consumer preference can be predicted, based on expert data. Minimum residues were given both during training and the test by the net consisting of 4 nodes. Validity of the neural network obtained was tested by predictions on random numbers, in addition to the interpretation of the residues. Based on the results thus obtained it was found that random numbers were predicted incorrectly by the network, no correlation was found between the data. In addition to the above, cross-checking of the model was also performed, based on which it showed no significant deviations. During the check, the net gave an accuracy value of 87%. based on this, the model was accepted, and the 4-node MLP model was used to predict the popularity of the other varieties, based on expert data.

According to the model developed, the most preferred variety on a 9-step scale was *‘Shinerock’* (8.46), while the least preferred varieties were predicted to be *‘Madonna’* and *‘Rustler’*, with average preference values of 2.7 (Table 4).

The order of importance of the variables that play a role in the structure of the nets was also given by the Palisade software during the training and testing of neural networks (Figure 2).

In the case of sweet corn varieties, during the search for the correlation between expert data and consumer preference data, the most important variables were found to be sweet taste (18%), global flavor intensity (14%) and juiciness (12%). This means that, based on the results of the network, higher preference scores were given by the consumers when evaluating juicy products with an intensive sweet flavor. Results support the conclusions of earlier studies [31], [32], [33].

Results thus obtained were then compared to the initial clusters from which the neural network was built, following the selection of the „most average” varieties (representing the members of the cluster the most). It can be seen from the results of **Table 5** that the sweet taste intensities of the first two clusters have significantly higher values than the other clusters (there is no clear separation from cluster 5).

Similarly, higher values were also received by the first two clusters for global flavor intensity, however, clusters 3 and 6 also have high values, so they differ significantly from clusters 4 and 5, based on the results of the Tukey-HSD test. It is important to note that, during the evaluation of global flavor intensity, panelists perform the evaluation on the basis of total flavor intensity, and these flavors do not necessarily represent an advantage during consumer evaluation. Based on tenderness, in addition to the first two clusters, members of cluster 6 possess significantly higher values compared to the other clusters, so the clusters tested were divided into two groups (**Table 6**).

Results show that the first two clusters possess intensive global and sweet taste, and these were significantly tender products. Results are also reflected in the predicted consumer results, because the first two clusters received almost identical results, an average preference value of 6.8, based on the neural network. The average preference value of the third cluster was 5.2, largely due to its high (80) global flavor intensity value. The next group consists of the fourth and sixth clusters with average preference values of 4.0 and 4.3, respectively. Members of the sixth cluster possess higher tenderness and global flavor intensity values, while members of the fourth cluster have a mediocre tenderness value. The least preferred samples are contained in the fifth cluster, with low intensity values for all three product properties important for preference prediction.

6. Conclusions

The presented approach, which combines the artificial neural networks presented and Monte Carlo simulation proved to be suitable to predict consumer preference, based on expert sensory testing results. The advantage of this approach is that consumer tests requiring a lot of time, energy and expenditure can be successfully replaced by predictions based on expert data. The preference values of 36 sweet corn samples were successfully predicted on the ba-

sis of the consumer evaluation of six samples. During the development of the artificial neural network model, product properties that are the main drivers of consumer acceptance were successfully identified. These are, in the order of importance: sweet taste, global flavor intensity and juiciness. In the future, it would be advisable to extend our research to other horticultural and food products, or use it in the case of products for which consumer preference is not influenced by only a few, easily identifiable product characteristics. Another option is the creation of a software implementation simplifying calculation steps, thus testing and validation of networks, as well as the prediction of new date could be performed by a single software.

In summary, it can be concluded that validated product-specific artificial neural networks can make the determination of the most important sensory properties possible. As a result of the new approach, results will be more reliable, repetitions can be carried out more easily, tests can be reproduced better, overall creating a time- and cost-effective analytical system.

7. Acknowledgement



Our research and work was carried out with the support of the ÚNKP-17-4 New National Excellence Program of the Ministry of Human Capacities, the János Bolyai research scholarship, and grant no. OTKA K112547. We express our thanks for the support.

8. Literature

- [1] McCulloch, W. S. and Pitts, W. H. (1943): A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115–137.
- [2] Hebb, D.O. (1949): *The organization of behavior: A neuropsychological theory*. New York, John Wiley and Sons. 335.
- [3] Rosenblatt, F. (1958): The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65 (6), 386–408.
- [4] Widrow, B. and Hoff, M.E. Jr. (1960): Adaptive Switching Circuits. IRE WESCON Convention Record. 96–104.
- [5] Borosy, A. P. (2001): Mesterséges ideghálózatok. In: Horvay szerk. *Sokváltozós adatelemzés (Kemometria)*. Budapest, Nemzeti tankönyvkiadó. 312–329.
- [6] Hopfield, J. (1982): Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci. U.S.A.*, 79 (8) 2554–2558.

- [7] Rumelhart, D.E. and McClelland, J.L., eds (1986) *Parallel Distributed Processing, Explorations in the Microstructure of Cognition*, MIT Press.
- [8] Horváth G. (szerk.) (2006): *Neurális hálózatok*. Altrichter M., Horváth G., Pataki B., Strausz Gy., Takács G., Vályon J., *Neurális hálózatok*, Budapest, Panem Kiadó, 2006.
- [9] Sipos László, Gere Attila, Kókai Zoltán, Szabó Dániel (2012): Mesterséges ideghálózatok (ANN) alkalmazása az érzékszervi minősítés gyakorlatában. 58, (1-2), 32-46.
- [10] Marini, F. (2009): Artificial neural networks in foodstuff analyses: Trends and perspectives. A review. *Analytica Chimica Acta*, 635, 121–131.
- [11] Debska B., Guzowska-Swider, B. (2011): Application of artificial neural network in food classification. *Analytica Chimica Acta*, 705, 283–291.
- [12] NeuralTools Version 5.7 Manual (2010).
- [13] Fu, LiMin (1994): *Neural Networks in Computer Intelligence*. McGraw-Hill, Inc. pp. 1-416.
- [14] Székely Géza, Losó Viktor, Tóth Arnold: Nemzetközi és hazai zöldség-gyümölcsfogyasztás, módszertani kérdések- ÉVIK 2015/1
- [15] Losó Viktor (2015): Gyorsfagyasztott csemegekukorica termékek komplex értékelése. *Doktori Értekezés*. Budapesti Corvinus Egyetem. 7-27.
- [16] Bíró, Gy., Lindner K. (Szerk.) (1999): *Tápanyagtáblázat: Táplálkozásban és tápanyagösszetétel*, Budapest, Medicina Kiadó.
- [17] Fodor, Z. (szerk.) (2016): *Zöldség és gyümölcs ágazat helyzete Magyarországon*. Budapest, FruitVeB Magyar Zöldség-Gyümölcs Szakmaközi Szervezet. 1-31.
- [18] Orosz, F. (2009): Termesztéstechnológiai elemek hatása acsemegekukorica koraiságára. *Doktori értekezés*, Budapest. p. 1- 157.
- [19] Kilcast, D. (2010): *Sensory analysis for food and beverage quality control*. Woodhead, Cambridge, UK.
- [20] MSZ ISO 6658:2007 *Érzékszervi vizsgálat. Módszertan. Általános útmutató*
- [21] Sipos, L. (2009): *Ásványvízfogyasztási szokások elemzése és ásványvizek érzékszervi vizsgálata*. Doktori értekezés. Budapesti Corvinus Egyetem. 80-102.
- [22] Kókai, Z. (2003): *Az almafajták érzékszervi bírálata*. Doktori értekezés. Budapest: Budapesti Közgazdaságtudományi és Államigazgatási Egyetem, 35-59.
- [23] Varela, P., Ares, G. (2014): *Novel Techniques in Sensory Characterization and Consumer Profiling*, CRC Press, pp. 9-41.
- [24] MSZ ISO 11035:2001 *Érzékszervi vizsgálat. A leíró kifejezések azonosítása és kiválasztása érzékszervi profilhoz többdimenziós eljárással*
- [25] ISO 13299:2003 *Sensory analysis – Methodology – General guidance for establishing a sensory profile*
- [26] MSZ EN ISO 8586:2014 *Érzékszervi vizsgálat. Általános útmutató a kiválasztott bírálók és az érzékszervi szakértő bírálók kiválasztásához, képzéséhez, valamint folyamatos ellenőrzéséhez*
- [27] MSZ EN ISO 8589:2015 *Érzékszervi vizsgálatok. Általános útmutató a bírálói helyiségek kialakításához*
- [28] Chen G.X, Jaradat S.A, Banerjee N., Tanaka T.S., Ko M.S.H., Zhang M.Q. (2002): Evaluation and comparison of clustering algorithms in analyzing ES cell gene expression data. *Stat Sin*, 12:241–262.
- [29] Héberger, K. (2010): Sum of ranking differences compares methods or models fairly. *Trend. Anal. Chem.* 29, 101–109.
- [30] Héberger, K., Kollár-Hunek, K. (2011): Sum of ranking differences for method discrimination and its validation: comparison of ranks with random numbers. *Journal of Chemometrics*, 25, (4) 151–158.
- [31] Gere, A., Losó, V., Tóth, A., Kókai, Z., Sipos, L. (2012): Kukorica fajták preferenciaterképezése szoftveres támogatással. *Élelmiszervizsgálati Közlemények*, 58, pp. 118-130.
- [32] Gere, A., Losó, V., Radványi, D., Juhász, R., Kókai, Z., Sipos, L. (2013): Csemegekukorica fajták komplex értékelése. *Élelmiszervizsgálati Közlemények*, 59, pp. 120-134.
- [33] Gere, A., Losó, V., Györey, A., Kovács, S., Huzsvai, L., Nábrádi, A. Kókai, Z., Sipos L. (2014): Applying parallel factor analysis and Tucker-3 methods on sensory and instrumental data to establish preference maps. Case study on sweet corn varieties. *Journal of the Science of Food and Agriculture*, 94, 15, pp. 3213-3225.