

Az orvosi Nobel-díjas Békésy György vegyész, fizikus emlékezete – 2022

Beszédalapú orvosi diagnosztika Békésy György nyomán

Vicsi Klára DSc

BESZÉDAKUSZTIKAI LABORATÓRIUM, BME VIK TMIT

BÉKÉSY EMLÉKÉV 2022

MTA Békésy György Akusztikai Kutatólaboratórium

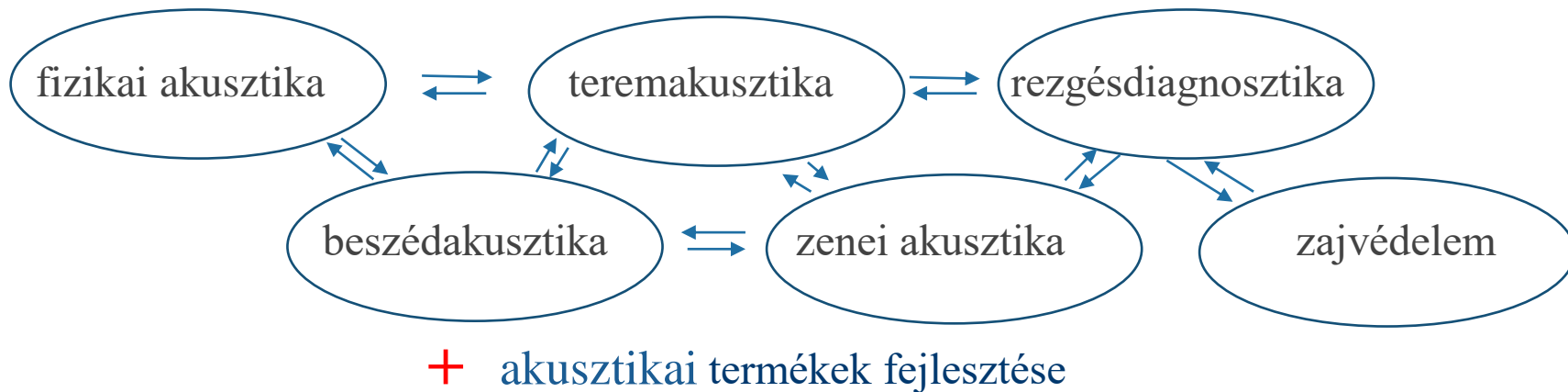
1975 őszén

Az MTA kutatási hálózatában európai viszonylatban is figyelemre méltó új független hazai akusztikai laboratórium jött létre, korszerű felszereltséggel Professor Tarnóczy Tamás (1975-1980), majd Dr. Illényi András vezetésével (1981-1995).

A Kutatólaboratórium az MTA Budaórsi úti Kutatóházában kapott helyet.



Kutatási területek 1975-1995



A laborban folyó kutatásra Békésy nyomában az **interdiszciplináris szemléletmód** volt jellemző.

Nemzetközi kapcsolatok kialakítása: Tarnóczy a Kelet-Európai Államok Akusztikai Szövetségének ötletadója és megszervezője (FASE). Ez később az Összeurópai Akusztikus Szövetség (EAA) egyik alapító szervezete lett.

A Kutatólaboratórium fő mérési helyiségei az akkor még Közép-Európában egyedülálló süketszoba és a zengőszoba.

Szabad hangtér akusztikai tulajdonságait
modellező **SÜKETSZOBA**
163m³ hasznos térfogatú



A testhang átvezetés csökkentése
érdekében a helyiségek alapja
rugókötegeken nyugszik



Diffúz hangteret modellező
ZENGŐSZOBA
131m³ térfogatú



A kiállításokra készült „korabeli” prospektus

1985-ben elkészült az első magyar kötött szótaras beszédfelismerő rendszer



MTA MŰSZERÜGYI ÉS
MÉRÉSTECHNIKAI SZOLGÁLAT

Békésy György Akusztikai Kutatólaboratórium

BESZÉDFELISMERŐ RENDSZEREK IBM XT/AT KOMPATIBILIS SZÁMÍTÓGÉPEKEN

Az MTA MMSZ Akusztikai Kutatólaboratóriumában kifejlesztett beszédfelismerő rendszerek lehetővé teszik, hogy személyi számítógépeken választott szavakkal és egyszerű mondatokkal számítógépeket, robotokat vagy egyéb berendezéseket lehessen vezérelni. Minden olyan területen igen hatékonyan alkalmazhatóak, ahol a berendezések szóval történő vezérése a kezek felszabadítását teszi szükségessé más tevékenység elvégzésére. Teszt fogyatékosok, betegek, akik kezükkel nem tudják a megszokott módon a környezetükben lévő berendezéseket irányítani, szóval történő vezérése alkalmazásával több hasznos tevékenység elvégzésére is képesek válnak.

Tipikus felhasználási területek:

GYÁRAK, ŰZEMEK:

- automaták, robotok, egyéb berendezések szóban történő irányítása
- Funkciók: adatgyűjtés, folyamatszabályozás, vezérlés, minőségellenőrzés

HIVATALOK:

- Billentyűzet részben vagy egészben történő kiváltása.

POSTA:

- CAD/CAM rendszereknél, térképésztesnél

KUTATÓ-FEJLESZTŐ INTÉZETEK:

- Levelek és csomagok szorításása

TESTI FOGYATÉKOSOK:

- magasszintű ember-gép kapcsolat

EGYEB:

- környezetben lévő berendezések vezérése, pl. telefon, rádió
- letárlás, megrendelésfelvétel, cím és üzenetregisztráció, stb.

Kiszótár kezelő rendszer (ST-03) /specifikáció/

Szótárkészlet: szabadon választható, változtatható szavak és tömdolok, nyelvtűggetten

Megengedett bemonási hossz: 2 s

Bemondások közötti minimális szünet: 0,5 s

Felismerési pontosság: 95-99% (szótármérettől és jel-zaj viszonytól függ; pontosság > 95% ha a szótárméret < 80 és a jel-zaj viszony > 20 dB)

Felismerési idő: < 500 ms

Többfelhasználós rendszer: felhasználók száma egyidőben max. 10

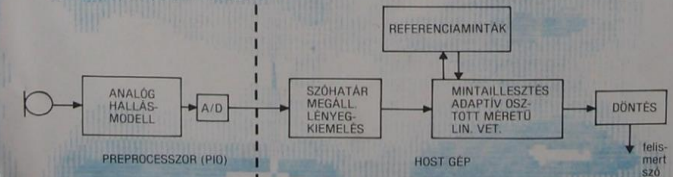
Előfeldolgozó egység: mikrofon bemenettel egy IBM kártyán

Felismerő rendszer: működik minden IBM XT/AT kompatibilis számítógépen

VEDELEM TEVESZTESEK ESETÉN!

A preprozessor a mikrofonnal felvett jel akusztikus elemzését 19 párhuzamos szűrőcsatornára alapuló hullámmódel segítségével végzi. A 10 ms időintervallumban mintavett paraméterek A/D átalakítás után a „párhuzamos porton” keresztül a HOST géphe kerülnek, ahol a további feldolgozás és a felismerés történik. A mintaillesztés adaptív időcsoportokban lineáris idővetéssel történik.

A RENSZER FELÉPÍTÉSE:



Rendszerváltás következményeként a Kutatólaboratórium átkerült a
Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Távközlési és Médiainformatikai Tanszékére.
A labor központi témája a beszédakusztika lett.
1996-ban a labor neve

BME TMIT Beszédakusztikai Laboratóriuma lett.

Kutatási területek 1996 -

Beszéd szöveggé alakítása

Jelentés feldolgozása (megértés)

Nyelv felismerése

Verbális tartalom feldolgozása

Beszélő személy felismerése, azonosítása

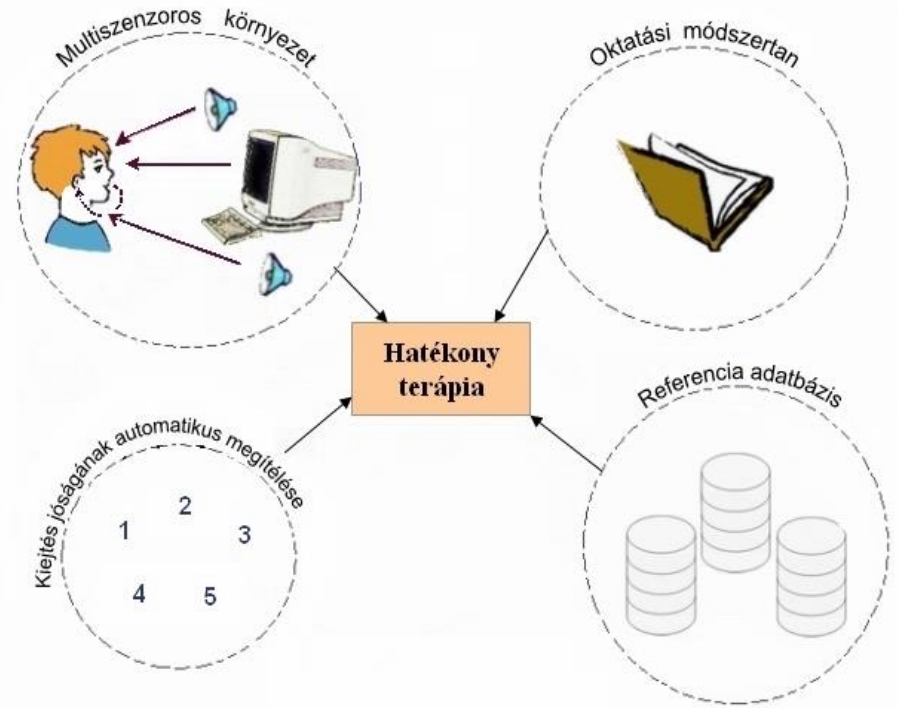
Nem-verbális tartalom feldolgozása

Érzelem, szándék, beszédstílus felismerése

Egészségi állapot felismerése



- Hallássérültek részére **audiovizuális oktatórendszert** hoztunk létre EU-s nemzetközi projekt keretében 5 nyelvre



- Hallástréning rendszereket** fejlesztettünk ki hallássérültek, cochleáris implantáltak részére.

A Békésy György kutatási hitvallásához híven nagy hangsúlyt fektettünk a mérnöki, fizikusi, nyelvészeti és az orvosi tudományterületen dolgozó kutatók összehangolt munkájára.

Az utóbbi 10 évben egyre inkább azon törtük a fejünket, hogy tudnánk az összegyűjtött tudásunkat az orvoslásban hasznosítani.

Hangdiagnosztika: beszéd, mint BIO-MARKER

**Hangdiagnosztikai eszközök fejlesztése,
automatikus diagnosztizálási lehetőségek**

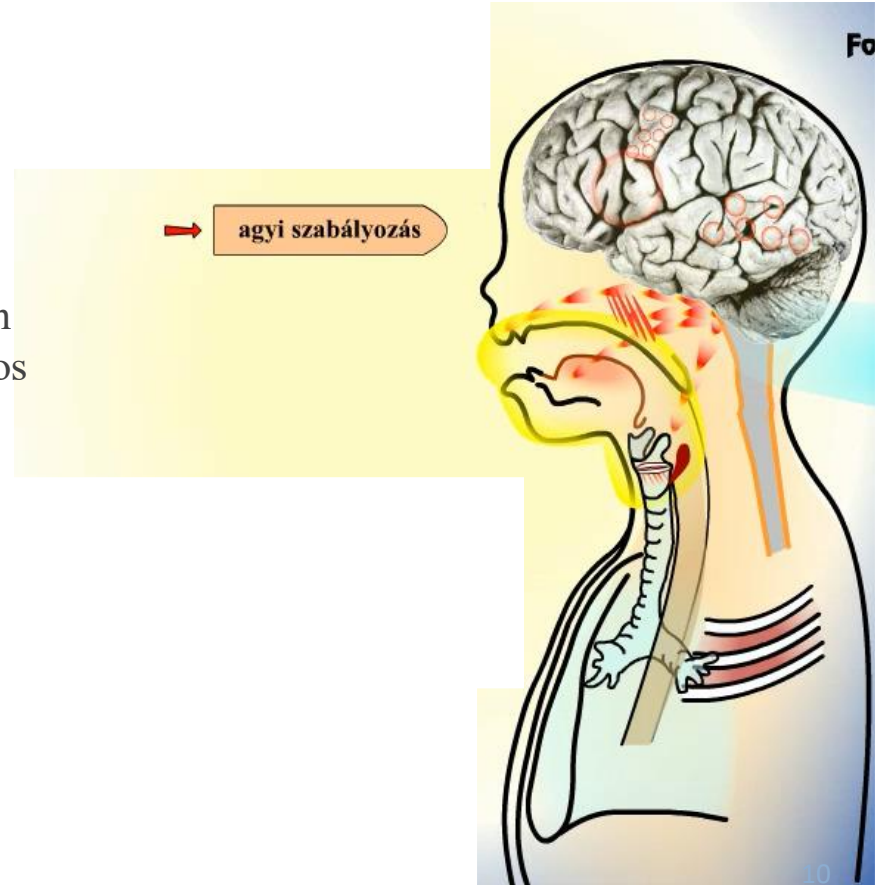
A beszéd, mint BIOMARKER

A beszéd létrehozása egy igen komplikált feladat, amelyhez számos agyterület és perifériás izomvezérlés összehangolt működése szükséges.

Ez az oka annak, hogy a hangképzőszervi elváltozásokon kívül, a beszéd akusztikai paraméterei érzékenyek számos neurológiai károsodásra is (cognitive dysfunctions).

A normálistól történő elváltozás:

- a hangszalag működés,
 - az artikuláció és a
 - prozódia
- területén is megjelennek, a probléma típusától függően.



A beszéd, mint BIOMARKER

Fő feladat: az adott betegségtípusra jellemző akusztikai paraméterek megkeresése

Ehhez szükséges:

```
graph LR; A[Ehhez szükséges:] --> B[A betegek hangját tartalmazó beszédadatbázis]; A --> C[Hasonló felépítésű egészséges beszédadatbázis];
```

A betegek hangját tartalmazó
beszédadatbázis

Hasonló felépítésű egészséges
beszédadatbázis

Hangdiagnosztikai eszközök fejlesztése

- Hogyan? \longrightarrow A statisztikai jelfeldolgozás eszközeivel.
- Elegendően sok minta alapján:
Jellemző beszédparaméterek valószínűségi eloszlásait vettük fel \longrightarrow probabilisztikus modellek
- Mesterséges intelligencia típusú osztályozás, regressziós eljárások, fuzzy alapú döntéstámogató rendszerek \longrightarrow gépi felismerés
- **A megvalósításhoz nagyméretű és jól feldolgozott, az adott probléma szempontjából optimalizált adatbázisokra van szükség!**



Hangképzőszervek fiziológiai elváltozásai:

- Hangképzőszervek elváltozásai és a beszéd bizonyos fizikai paramétereinek között szoros összefüggés van.
 - Pl. a gégefő sérülésekor a hangszalag rezgése a normálistól eltérő lesz.



Gégedaganat

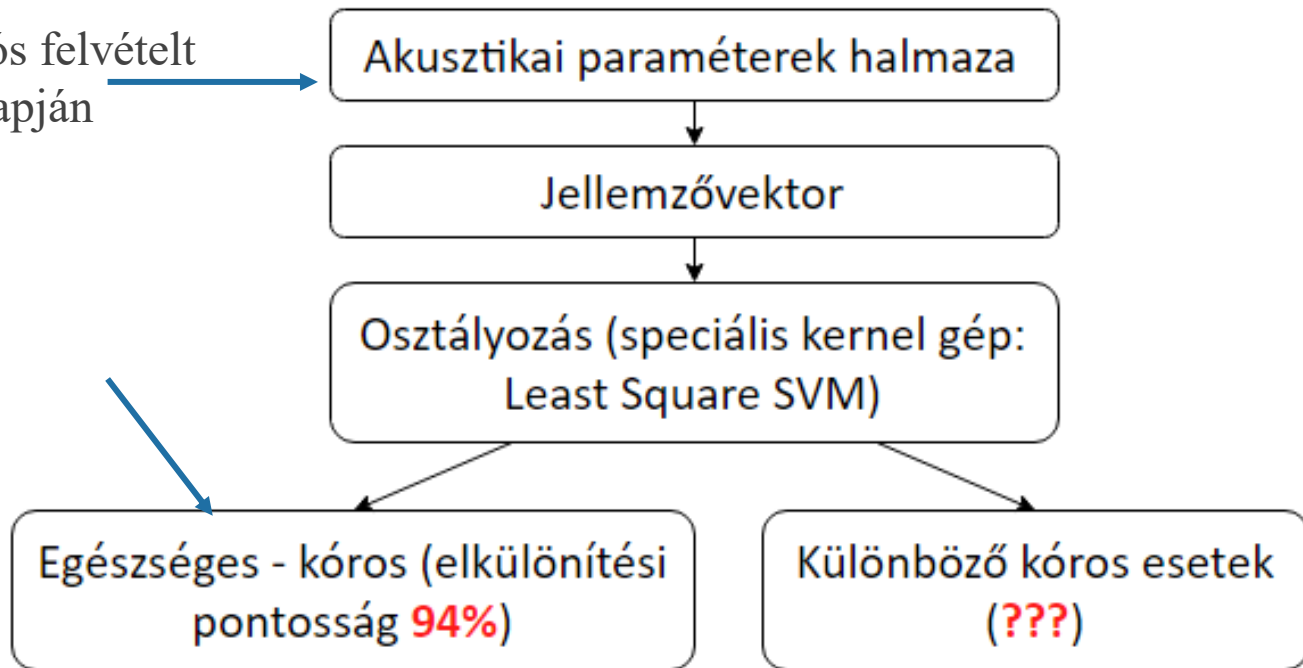


Hangszalagpolip

Gégészeti diagnosztikai rendszer fejlesztése – egészséges - kóros megkülönböztetés

Az Országos Onkológiai Intézet Fej és Nyaksebészeti Osztályával együttműködésben

119 egészséges és 152 kóros felvételt
tartalmazó mintahalmaz alapján



Pszichológiai, neurológiai elváltozások

- A beszéd **különböző agyi területek, és periférikus izomkontrollok koordinációját igényli!**
- Ez okozza a beszéd fizikai paramétereinek a neurológiai defektusokra (kognitív diszfunkciók) való nagyfokú érzékenységét.
- A hiányosságok hangképzési, artikulációs, prozódiai területen, és a beszéd folyamatosságában jelennek meg.
- Parkinson kór, demencia, agyvérzés, depresszió, hipoxia, stb.

Depresszió - pszichiátriai megbetegedés

Olyan hosszan tartó elmeállapot, aminek jellemző tünetei: lehangoltság, szomorúság, veszteségérzet, elesettség. A pszichotikus depressziót gyakran kíséri halálvágy, öngyilkossági késztetések.



Hatása a beszédre

az agy motorikus működését befolyásolja, emiatt változik a depressziós ember beszédproduktuma.

- **Orvosok jelzői:** fakó, monoton, élettelen, lomha, és fémes stb.



Depresszió – egészséges –kóros megkülönböztetés

Semmelweis Orvostudományi Egyetem, Pszichiátriai Klinikával együttműködésben

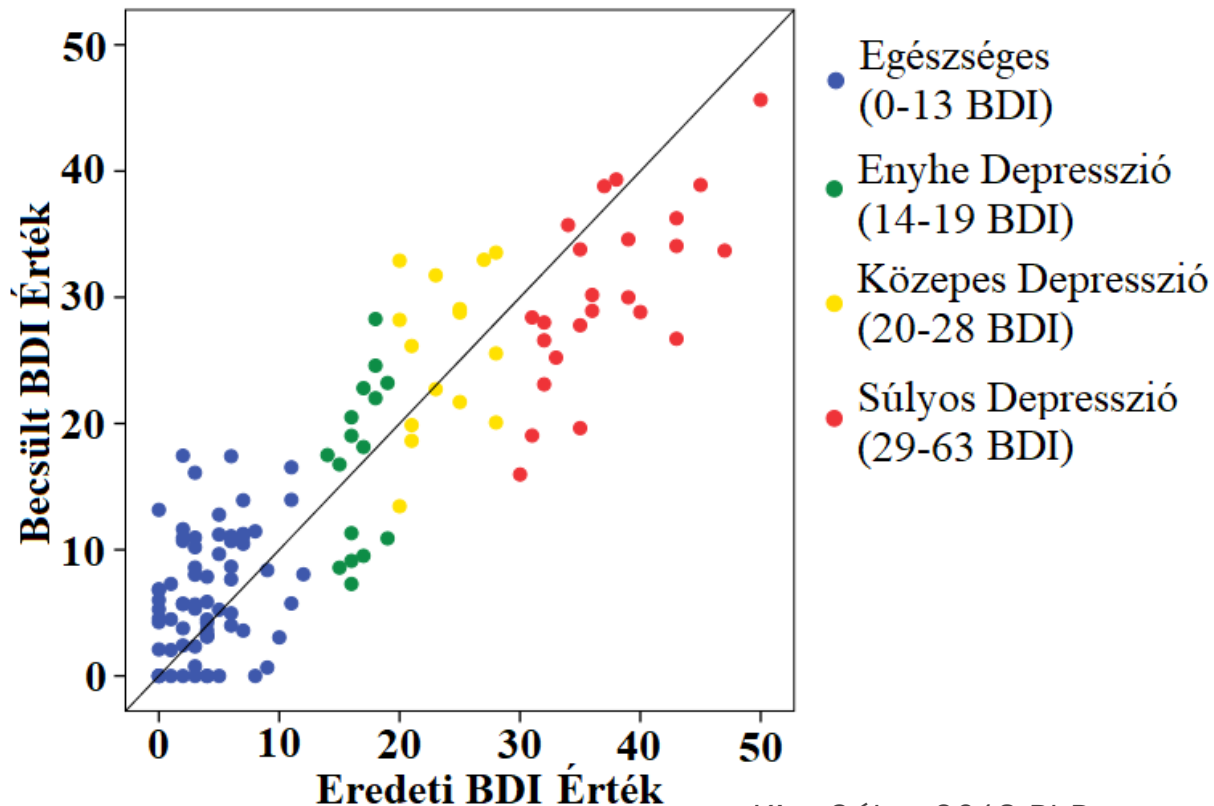
Betanítás: 55 depressziós és egészséges hanganyag
22 paraméteres jellemzővektor,

Osztályozó: szupport vektor gép (SVM)

Elkülönítési pontosság: **85%**

Depresszió súlyosságának becslése

BDI értékek becslése:
Szuport Vektor Regressziós
eljárással



Pszichológiai állapotkövető rendszer fejlesztése



- Pszichológiai, pszichiátriai, nyelvész, mérnök kutatók nemzetközi együttműködésében, az Európai Űrkutatás Ügynökség megbízásából
- szezonális depresszió automatikus detektálására
- CONCORDIA Űrállomás a Déli Sarkon

*COALA - Psychological Status Monitoring by Computerised Analysis of Language Phenomena ESA
PECS 4000108003/13/NL/KML*

Concordia kutatóállomás,
Antarktika, Dome C (2005)

3 233 m tengerszint fölött,
a geológiai déli sarktól
1670 kilométerre

Dome C a Föld leghidegebb
pontja: -25 °C nyáron, -80 °C
télen ($-84,6\text{ °C}$ 2010-ben).
Átlaghőmérséklet: $-54,5\text{ °C}$

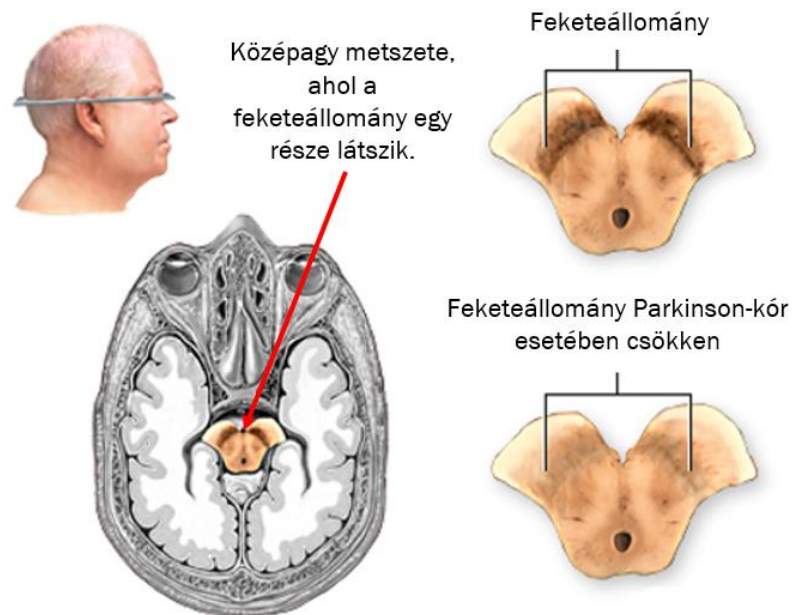
Páratartalom alacsony, száraz,
kevés a csapadék. Átlagos
légnyomás érték 645 hPa .

Sok a hasonlóság az űrrel
(mély-űr kutatás, pl. 9 hónapig
sötétség van).



Parkinson Kór - degeneratív idegrendszeri betegség

- Az agyban a finom mozgásokat irányító bazális ganglionok mennek tönkre, és csökken a fő ingerületátvivő anyag, a dopamin termelődése is.
- Tünet: a nyugalmi remegés, izommerevség és a meglassultság
- Nem megfordítható, de szinten tartható betegség.



Parkinson Kór

Hatása a beszédre: halk beszéd, rendellenes hangszalag remegés, ziháló hangadás, nehéz hangindítás, akadozó beszéd.



Akusztikai paraméterek  Komplex hatás:

- zöngképzést befolyásoló paraméterek, mint pl. Jitter, Shimmer,
- alapfrekvencia változásához kapcsolódó paraméterek,
- Egy-egy magánhangzó intenzitásváltozására jellemző paraméterek,
- artikulációhoz kapcsolódó paraméterek, mint a hangképző üregek rezonancia adatai,
- tranziens és stacionárius részek aránya a folyamatos beszédben,
- időtartam arányokhoz tartozó paraméterek.

Parkinson-kór : egészséges – kóros megkülönböztetés

Semmelweis Orvostudományi Egyetem és a Szt. János Kórház idegsebészetével együttműködésben

Bemenet: 20 paraméteres jellemzővektor

Osztályozó: szupport vektor gép (SVM)

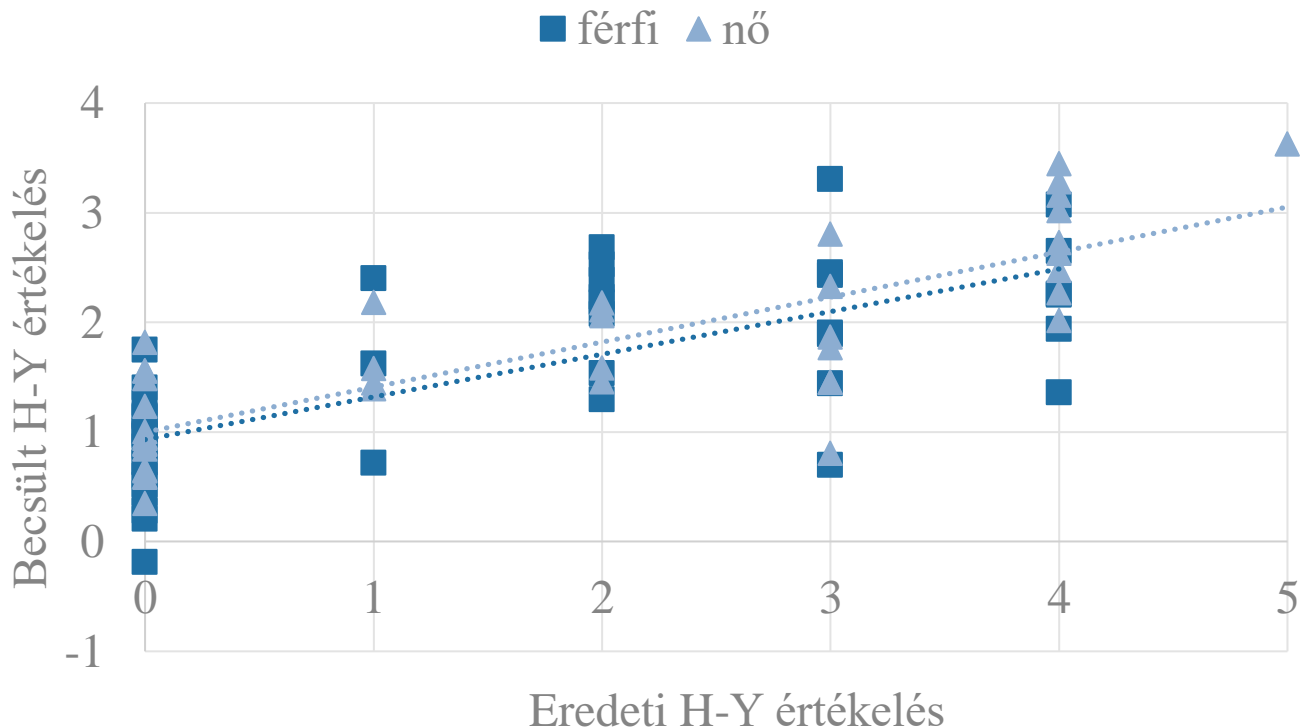
Betanítás: 55 beteg hangfelvétel

Elkülönítési pontosság: **88 %**

Parkinson kór súlyosságának becslése

20 paraméteres jellemzővektor, Lineáris regresszió, betanítás 28 férfi és 27 női hangfelvétellel

A becslés átlagos négyzetes hibája:
1,01 H-Y



A betegségek együttes diagnosztizálása

Hangképzőszervi megbetegedés, depresszió, Parkinson-kór, egészséges csoport

Hanganyag: 5 mondatból álló mese felolvasása

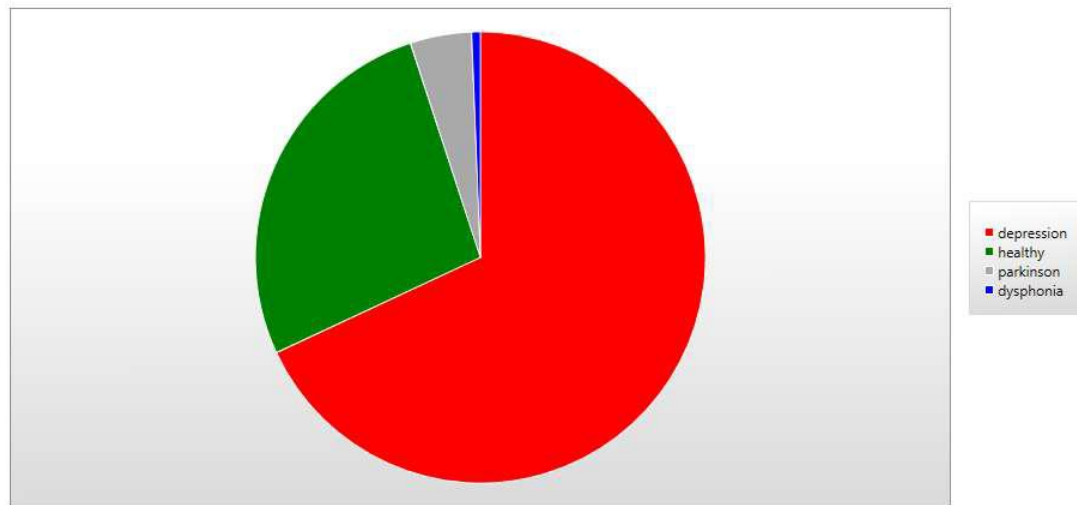
Bemenet: 270 paraméteres jellemzővektor, amelyből 30 lett kiválasztva

Osztályozó: szupport vektor gép
(SVM-rbf)

Elkülönítési pontosság:

4 osztályos megkülönböztetés: **88 %**

Hierarchikus osztályozás **86%**



healthy: 26.9%

depression: 68.1%

dysphonia: 0.6%

parkinson: 4.4%

Köszönöm a figyelmet!

vicsi.klara@vik.bme.hu

Békésy emlékévé 2022

