

Analiza glasu kot diagnostična metoda za odkrivanje Parkinsonove bolezni

Speech Analysis as a Diagnostic Method for the Detection of Parkinson's Disease

Andraž Levstek
Gimnazija Jožeta Plečnika
Ljubljana, Slovenija
levstek.andraz@gmail.com

Darja Silan
Gimnazija Jožeta Plečnika
Ljubljana, Slovenija
darja.silan@gjp.si

Aljoša Vodopija
Institut "Jožef Stefan"
Ljubljana, Slovenija
aljosa.vodopija@ijs.si

POVZETEK

Parkinsonova bolezen je nevrodegenerativna bolezen, ki povzroča težave v delovanju mišic zaradi pomanjkanja dopamina v možganskem deblu. Poleg na delovanje mišic vpliva Parkinsonova bolezen tudi na glas. Slednji postane bolj monoton, hripav in šibek. Zaradi naštetih sprememb se za diagnosticiranje Parkinsonove bolezni vse pogosteje uporablja analiza glasu z metodami umetne inteligence. V tej raziskavi s pomočjo metod strojnega učenja primerjamo zvočne posnetke glasu zdravih oseb in bolnikov s Parkinsonovo boleznijo. Za izboljšavo klasifikacijske točnosti se dodatno poslužujemo pristopu zmanjševanja razsežnosti. Najbolj točen klasifikator zgradimo z uporabo metode naključnih gozdov, s katerim dosežemo 73 % točnost. S tem potrdimo povezavo med Parkinsonovo boleznijo in karakteristično spremembo glasu. V nadaljevanju ocenimo pomembnost posameznih zvočnih posnetkov in pripadajočih atributov. Izsledke raziskave lahko uporabimo za nadgradnjo obstoječe metodologije s predlogi za dodatne zvočne posnetke, ki bolje posnemajo obstoječe najboljše posnetke glasu.

ABSTRACT

Parkinson's disease is a neurodegenerative disease, that causes impaired muscle function, because of a lack of dopamine in the brain stem. Parkinson's disease also affects speech ability. The voice becomes monotone, hoarse and feeble. For this reason, one of the emerging ways to diagnose Parkinson's disease is becoming the speech analysis by applying artificial intelligence. In this paper, we use machine learning to connect voice samples to the presence of Parkinson's disease. To improve the classification accuracy, we additionally use a dimensionality reduction approach. The most accurate classifier was built with random forest, with an accuracy of 73 %. The experimental results confirm the link between the voice changes and the presence of Parkinson's disease. Additionally, we estimate the importance of individual voice samples and corresponding features. The results can be used to improve the current speech analysis by proposing additional voice samples, that more accurately reflect the properties of the current best samples.

KLJUČNE BESEDE

Parkinsonova bolezen, analiza glasu, strojno učenje, naključni gozdovi, pomembnost atributov

KEYWORDS

Parkinson's disease, speech analysis, machine learning, random forest, feature importance

1 UVOD

Parkinsonova bolezen je nevrodegenerativno in izčrpavajoče bolezensko stanje, ki vpliva na osrednje živčevje. Bolezen prizadene približno 1 % ljudi starejših od 60 let. Bolnik s Parkinsonovo boleznijo se pogosto tresi, ima težave s hojo in ravnotežjem, njegovo gibanje postane počasno, pojavi se rigidnost. Pojavijo se lahko tudi duševne motnje, kot so anksioznost, depresija ter težave s spanjem, razmišljanjem in obnašanjem.

Parkinsonova bolezen vpliva tudi na glas. Večina bolnikov ima govorne težave, kot so šibek, zadihan, hripav, višji in monoton glas. Za bolnika je značilna hripavost, zmanjšana jakost glasu, težava s pravilno artikulacijo fonemov in brbljanje. Bolnik težje najde prave besede in težje sodeluje pri hitrih pogovorih [5].

Diagnostično metodo, ki bi stoodstotno dokazala prisotnost Parkinsonove bolezni, še ne poznamo. Diagnoza temelji na vidnih in razpoznavnih simptomih, preteklem zdravstvenem stanju, fizičnem ter nevrološkem pregledu in bolnikovi anamnezi [13]. Po kriterijih mora biti za dokaz Parkinsonove bolezni prisotna akineza ter še vsaj ena druga lastnost (npr. tremor rok pri mirovanju, rigidnost ter posturalne motnje). Po teh kriterijih se Parkinsonovo bolezen lahko identificira z 90 % točnostjo, vendar diagnoza traja tudi več let [12]. Pri diagnosticiranju se uporablja tudi slikanje možganov z magnetno resonanco, pozitronsko emisijsko tomografijo in računalniško tomografijo. Vse našete diagnostične metode so drage ter zahtevne, zato se išče cenejše in preprostejše metode [13].

V diagnostične namene se vse pogosteje uporablja analiza zvočnih posnetkov glasu z uporabo metod umetne inteligence (npr. strojno učenje, procesiranje signalov itd.). Tovrsten način diagnostike je povsem varen, preprost, hiter in ne zahteva dragocenih namenskih naprav [8]. Vendar je to področje v primeru Parkinsonove bolezni še v razvoju. Večina raziskovalcev se namreč ukvarja le z doseganjem čim večje klasifikacijske točnosti [1, 7, 10, 11]. Pri tem pa zanemarjajo pomemben aspekt analize, in sicer da bi skušali identificirati pomembne posnetke in pripadajoče glasovne attribute. Slednja dognanja bi pripomogla k boljšem razumevanju problematike in omogočila oblikovanje bolj natančnih testov.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Information society '20, October 5–9, 2020, Ljubljana, Slovenia

© 2020 Copyright held by the owner/author(s).

V tem prispevku testiramo uporabnost analize glasu z metodami strojnega učenja za diagnosticiranje Parkinsonove bolezni. Opravljena študija temelji na zvočnih posnetkih 40 oseb (20 bolnikov s Parkinsonovo boleznijo) pridobljenih v raziskavi [10]. Na teh podatkih testiramo pet različnih algoritmov strojnega učenja. Za izboljšanje rezultatov dodatno uporabimo metodo za zmanjšanje razsežnosti in izboljšamo klasifikacijsko točnost za približno 5 %.

V nasprotju z večino sorodnega dela ocenimo tudi pomembnost posameznih posnetkov in pripadajočih atributov. V ta namen uporabimo metodo naključnih gozdov, saj ta dosega najvišjo točnost. Na ta način lahko ugotovimo kateri atributi in posnetki vsebujejo več informacij o prisotnosti Parkinsonove bolezni.

Prispevek je organiziran na sledeči način. V drugem poglavju predstavimo podatke. V tretjem poglavju opišemo metodologijo. V četrtem in petem poglavju predstavimo rezultate in pridobljena dognanja. V zadnjem poglavju naredimo zaključek in orisemo nadaljnje delo.

2 PODATKI

Podatki so bili zbrani na Istanbulski fakulteti za medicino (Istanbul Faculty of Medicine, Istanbul University) leta 2014 [10]. Zbrali so zvočne posnetke 40 ljudi, 6 žensk ter 14 moških s Parkinsonovo boleznijo in 10 zdravih žensk ter 10 zdravih moških. Vsaka oseba je podala 26 posnetkov, ki vključujejo samoglasnike, kratke stavke in besede. Natančneje, posnetki 1–3 predstavljajo trajajoče samoglasnike "a", "o" in "u", posnetki 4–13 števila od 1 do 10, posnetki 14–17 krajše stavke in posnetki 18–26 besede. Vsi zvočni posnetki so v turščini.

Vsaki osebi pripada 26 zvočnih posnetkov in vsakemu posnetku 26 linearnih ter frekvenčnih atributov zgrajenih z uporabo programske opreme za akustično analizo Praat [2]. Vsi atributi so numerični in se jih običajno izračuna za analizo glasu [2, 10]. Povzeti so v Tabeli 1. Skupno je v množici podatkov tako 676 atributov in ciljni razred. Slednji je binaren in predstavlja prisotnost (pozitiven = 1) oziroma odsotnost (negativen = 0) Parkinsonove bolezni. Imena nekaterih atributov uporabljamo v angleščini, saj pripadajoči slovenski izraz ne obstaja.

3 METODOLOGIJA

Klasifikatorje smo gradili s petimi algoritmi za strojno učenje: odločitveno drevo (C4.5), naivni Bayes (NB), metoda najbližjih sosedov (*k*NN), metoda podpornih vektorjev (SVM) ter metoda naključnih gozdov (RF). Za vse navedene algoritme smo uporabili privzete vrednosti parametrov, saj uglaševanje ni signifikantno izboljšalo klasifikacijske točnosti.

Število atributov močno presega število primerkov, zato smo se odločili za uporabo metod zmanjševanja razsežnosti in s tem uspešno izboljšali klasifikacijsko točnost za 5 %. Za izbor atributov smo uporabili široko poznano metodo imenovano rekurzivna odstranitev atributov (ang. *recursive feature elimination*, RFE) [4], ki temelji na vzvratni odstranitvi nepomembnih atributov. RFE spada med metode po principu ovojnice (ang. *wrapper*) in jo lahko uporabimo v kombinaciji z zgoraj naštetimi algoritmi za strojno učenje.

Za strojno učenje smo uporabili knjižnico *caret* [6] implementirano v programskem jeziku R [9].

4 REZULTATI

Za namen evaluacije smo uporabili pristop po metodi "izpusti enega" (ang. *leave one subject out*, LOSO). Najprej smo na učni

Tabela 1: Glasovni atributi uporabljeni za strojno učenje: frekvenčni, pulzni, amplitudni, glasovni ter harmonični.

Skupina	Atribut
Frekvenčni	Jitter (local)
	Jitter (local, absolute)
	Jitter (rap)
	Jitter (ppq5)
	Jitter (ddp)
Pulzni	Število glasovnih pulzov
	Število nihalnih dob
	Povprečna perioda
	Standardna deviacija period
Amplitudni	Shimmer (local)
	Shimmer (local)
	Shimmer (local, dB)
	Shimmer (apq3)
	Shimmer (apq5)
	Shimmer (apq11)
Glasovni	Delež nezvenečih časovnih oken
	Število lomljen glasu
	Delež lomljenj glasu
Harmonični	Srednja vrednost višine glasu
	Povprečna višina glasu
	Standardna deviacija višine glasu
	Najvišja višina tona
	Najnižja višina tona
	Avtokorelacija tona
Razmerje šum-harmonik	
	Razmerje harmonik-šum

Tabela 2: Rezultati klasifikatorjev v obliki točnosti, senzitivnosti in specifičnosti. Najvišja vrednost posamezne metrike je odebena.

Algoritem	Točnost	Senzitivnost	Specifičnost
C4.5	0,63	0,65	0,60
NB	0,63	0,80	0,45
<i>k</i> NN	0,48	0,55	0,40
SVM	0,68	0,70	0,65
RF	0,73	0,75	0,70

Tabela 3: Matrika zamenjav za klasifikator zgrajen z metodo RF.

Napoved / Pravi	Negativen (0)	Pozitiven (1)
Negativen (0)	14	5
Pozitiven (1)	6	15

množici uporabili metodo RFE za izbor najboljših atributov ter za izbrane attribute zgradili klasifikator. Nato smo klasificirali izpuščen primerek. Postopek smo ponovili za vse primerke.

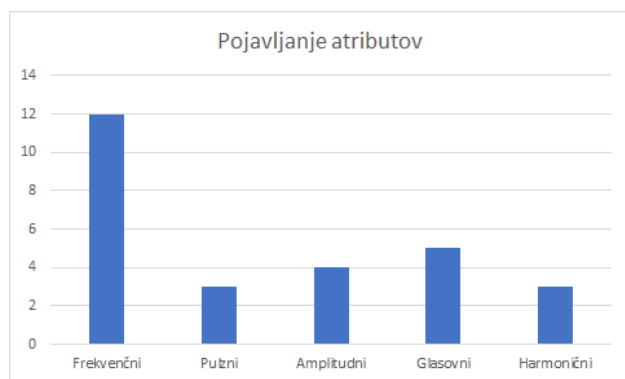
V Tabeli 2 so prikazani rezultati v obliki povprečne točnosti, povprečne senzitivnosti in povprečne specifičnosti. Vidimo, da je najbolj točen klasifikator zgrajen z metodo RF, najmanj točen pa z metodo *k*NN. Najvišjo senzitivnost je dosegel klasifikator zgrajen z metodo NB, specifičnost pa z metodo RF. V Tabeli 3 so prikazani

rezultati za klasifikator zgrajen z metodo RF v obliki matrike zamenjav. Klasifikator je pravilno klasificirala 29 primerov, zmotil pa se je v 11 primerih.

Na koncu smo na celotnih podatkih uporabili kombinacijo izbora atributov po metodi RFE in algoritma za strojno učenje RF, saj je le ta dosegel najvišjo točnost. Zanimalo nas je pojavljanje posameznih posnetkov in pripadajočih atributov ter pomembnost obeh. Pomembnost izbranih posnetkov in atributov smo izračunali z metodo imenovana permutacijska pomembnost (ang. *permutation importance*), ki se jo neposredno izračuna z metodo RF [3].

Metoda RFE izbere 27 izmed 676 atributov. Med njimi se najpogosteje pojavljajo frekvenčni atributi (Slika 1). Medtem ko so ostale skupine atributov podobno zastopane. Med posnetki se najpogosteje pojavljajo števila, nato kratki stavki. Najslabše zastopani so trajajoči samoglasniki (Slika 2).

Slika 3 in Slika 4 predstavljata zaporedoma pomembnost atributov in posnetkov za metodo RF. Razvrščeni so od manj pomembnih do bolj pomembnih. Iz rezultatov je razvidno, da so frekvenčni atributi najbolj pomembni za metodo RF. Najmanj pomembni pa so harmonični atributi in atributi izpeljani iz tona glasu. Najpomembnejši posnetek je število "4". Opazimo, da števila in kratki stavki vsebujejo več informacij od ostalih posnetkov.



Slika 1: Pojavljanje skupin atributov (frekvenčni, pulzni, amplitudni, glasovni ter harmonični) po izboru z metodo RFE v kombinaciji z RF.



Slika 2: Pojavljanje skupin posnetkov (trajajoči samoglasniki, števila, kratki stavki ter besede) po izboru z metodo RFE v kombinaciji z RF.

5 DISKUSIJA

Podobno kot sorodne raziskave [1, 7, 10, 11] tudi naši rezultati dokazujejo povezavo med glasovnimi atributi in prisotnostjo (oziroma odsotnostjo) Parkinsonove bolezni. Z metodo RF, ki zgradi najbolj točen klasifikator, uspešno klasificiramo 73 % primerkov. Pri tem so najbolj pomembni in pogosti frekvenčni atribut (Slika 1 in Slika 3). Sklepamo, da zaradi deviacij frekvence glasu pri Parkinsonovi bolezni. Med posnetki izstopajo števila in kratki stavki (Slika 2 in Slika 4). O prisotnosti bolezni povedo več zahtevni ter daljši posnetki.

Ne glede na to, je tak način diagnoze nezadosten. Najbolj točna metoda zgreši 25 % bolnikov, kar je za medicinsko prakso nesprejemljivo [13]. Pri tem moramo poudariti, da smo imeli opravka z omejenim številom primerkov (posnetih je bilo le 40 oseb). V primeru, da bi zbrali več zvočnih posnetkov tako obolelih kot zdravih oseb, bi lahko klasifikator izboljšali z uporabo naprednejših metod strojnega učenja, ki jih na tako malem številu primerkov ni bilo moč uporabiti.

Morda ne bo nikoli možno določiti prisotnost Parkinsonove bolezni iz analize glasu z uporabo metod strojnega učenja, vendar bi tovrstne metode lahko uporabili bodisi komplementarno za nadgradnjo obstoječih metod bodisi kot presejalni test. Pri tem poudarimo, da je analiza glasu poceni in za bolnika povsem nemoteča ter varna preiskava.

6 ZAKLJUČEK

V prispevku smo z metodami strojnega učenja primerjali zvočne posnetke zdravih oseb in bolnikov s Parkinsonovo boleznijo. Namen študije je bil preveriti, ali lahko iz analize glasu sklepamo o prisotnosti Parkinsonove bolezni in ali je možno zgraditi klasifikator za uporabo v praksi. Dodatno smo tudi ocenili pomembnost posameznih posnetkov in pripadajočih glasovnih atributov.

Rezultati nakazujejo, da pri bolnikih s Parkinsonovo boleznijo pride do poslabšanja zvočne artikulacije, saj smo s klasifikatorjem zgrajenim z metodo naključnih gozdov uspešno zaznali 73 % bolnikov. Ne glede na to, klasifikator še ni primeren za uporabo v praksi, saj je njegova točnost prenizka. Sedanji klasifikator lahko uporabimo kot komplementarni test že obstoječim.

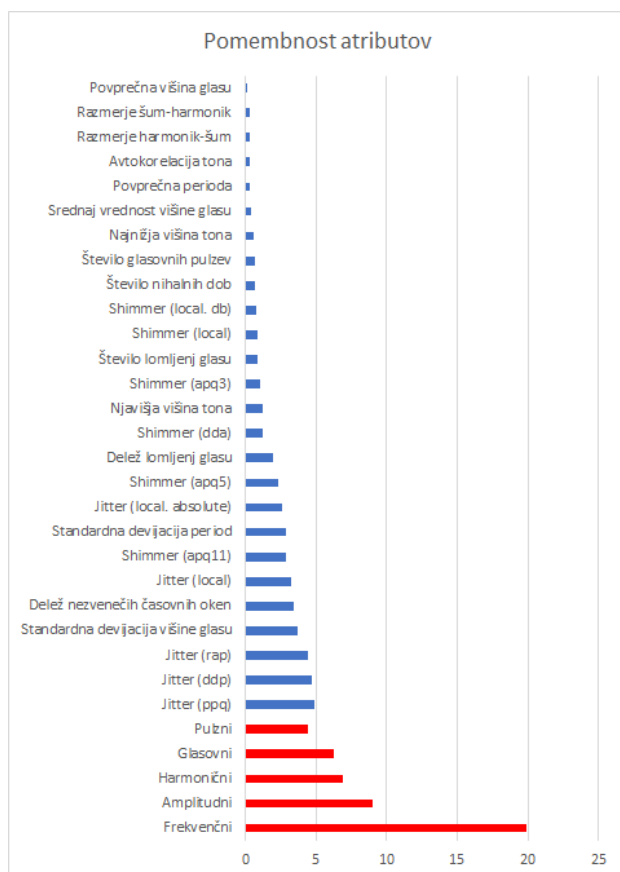
Za najbolj pomembne zvočne posnetke se izkažejo števila in kratki stavki. Pri tem so najmanj pomembni trajajoči samoglasniki in besede. Med atributi izstopajo frekvenčni in amplitudni.

Trenutno raziskujemo možnost, da bi zbrali več sorodnih zvočnih posnetkov. Na ta način bi lahko uporabili kompleksnejše metode strojnega učenja, ki omogočajo odkrivanje zagonetnih zakonitosti, ki jih na tako majhnem naboru primerkov ni bilo moč odkriti.

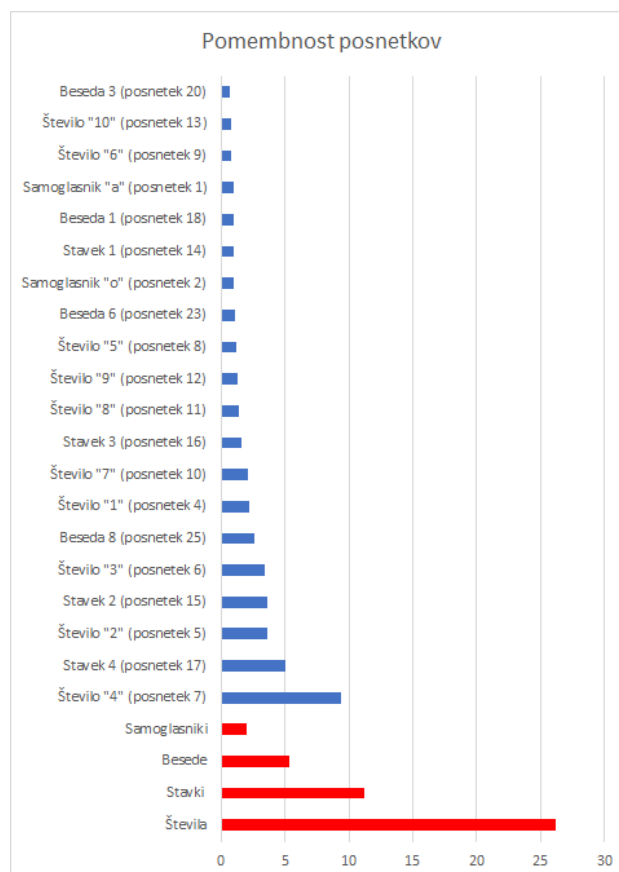
Naš dolgoročni cilj je izgradnja klasifikatorja, ki bi uspešno identificiral večino bolnikov tudi za ceno nekoliko slabše točnosti (nekateri zdrave osebe klasificiral za bolne). Tovrsten klasifikator, bi nato lahko uporabili kot presejalni test in na ta način olajšali sedanjo diagnostiko Parkinsonove bolezni. Poskusili bomo tudi razbrati zakaj so posnetki števil vsebovali več informacij o prisotnosti bolezni in z dobljenim znanjem skušali predlagati boljše celovitejši nabor izrazov, besed in fonemov.

ZAHVALA

Avtorji se zahvaljujejo gospe Ireni Hočevar Boltežar za razlago glasovnih atributov in slovenske prevode. A. Vodopija se dodatno zahvaljuje finančni podpori Javni agenciji za raziskovalno dejavnost Republike Slovenije (program usposabljanja mladega raziskovalca).



Slika 3: Pomembnost izbranih atributov za klasifikator zgrajen z metodo RF.



Slika 4: Pomembnost izbranih posnetkov za klasifikator zgrajen z metodo RF.

LITERATURA

- [1] I. Bhattacharya in M. P. S. Bhatia. 2010. SVM classification to distinguish parkinson disease patients. V *Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India*. ACM, New York, NY, USA, 1–6. doi: 10.1145/1858378.1858392.
- [2] P. Boersma. 2001. Praat, a system for doing phonetics by computer. *Glott International*, 5, 9/10, 341–345.
- [3] L. Breiman. 2001. Random forests. *Machine Learning*, 45, 1, 5–32. doi: 10.1023/A:1010933404324.
- [4] I. Guyon, J. Weston, S. Barnhill in V. Vapnik. 2002. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46, 1, 389–422. doi: 10.1023/A:1012487302797.
- [5] I. Hočvar Boltežar. 2013. *Fiziologija in patologija glasu ter izbrana poglavja iz patologije govora*. Pedagoška fakulteta. <http://www.biblos.si/lib/book/9789612531416>.
- [6] M. Kuhn. 2008. Building predictive models in R using the caret package. *Journal of Statistical Software, Articles*, 28, 5, 1–26. doi: 10.18637/jss.v028.i05.
- [7] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman in L. O. Ramig. 2009. Suitability of dysphonia measurements for telemonitoring of parkinson's disease. *IEEE Transactions on Biomedical Engineering*, 56, 4, 1015–1022. doi: 10.1109/TBME.2008.2005954.
- [8] M. A. Little, P. E. McSharry, S. Roberts, D. Costello in I. Moroz. 2007. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *Nature Precedings*. doi: 10.1038/npre.2007.326.1.
- [9] R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <http://www.R-project.org/>.
- [10] B. E. Sakar, M. E. Isenkul, C. O. Sakar, A. Sertbas, F. Gurgun, S. Delil, H. Apaydin in O. Kursun. 2013. Collection and analysis of a parkinson speech dataset with multiple types of sound recordings. *IEEE Journal of Biomedical and Health Informatics*, 17, 4, 828–834. doi: 10.1109/JBHI.2013.2245674.
- [11] C. O. Sakar in O. Kursun. 2010. Tediagnosis of parkinson's disease using measurements of dysphonia. *Journal of Medical Systems*, 34, 4, 591–599. doi: 10.1007/s10916-009-9272-y.
- [12] C. Silva. 2018. Speech analysis may help diagnose parkinson's and at earlier stage, study says. *Parkinson's News Today*. (2018). <https://parkinsonsnewstoday.com/2018/02/05/speech-analysis-can-help-detect-parkinsons-in-early-stages-study-says/>.
- [13] E. Tolosa, G. Wenning in W. Poewe. 2006. The diagnosis of parkinson's disease. *The Lancet Neurology*, 5, 1, 75–86. doi: 10.1016/S1474-4422(05)70285-4.