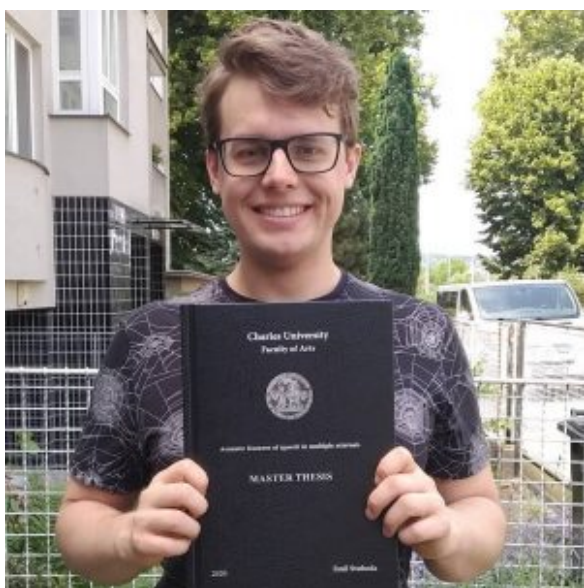


Small Talk: Roztroušená skleróza mění lidský hlas. Fonetik pracuje na rychlejším odhalení nemoci

Že z hlasu druhého člověka dokážeme vyčíst jeho emocionální rozpoložení, je vcelku nasnadě. Že se v něm ale odráží i jeho zdravotní stav, je věc poměrně nová a neprozkoumaná. V současné době se jí věnují například vědci z ČVUT, ale i absolvent Fonetického ústavu FF UK a doktorand Ústavu formální a aplikované lingvistiky MFF UK Emil Svoboda.

Výzkum, který Emil Svoboda poprvé představil ve své diplomové práci a na němž pracuje dodnes, se soustředí na rozpoznávání symptomů roztroušené sklerózy (RS) v hlase. Analýza jemných a opakujících se akustických odchylek dokáže s pomocí strojového učení urychlit diagnostiku nemoci.



Fonetik Emil Svoboda se svou diplomovou prací, na kterou navazuje další slibný výzkum. Zdroj: Archiv ES

Ja
ko
ne
po
uč
en
ý
la
ik
,
kt
er
ý
o
ro
zt

ro
uš
en
é
sk
le
ró
ze
sl
yš
el
po
dr
ob
ně
ji
na
po
sl
ed
y
na
st
ře
dn
í
šk
ol
e,
js
em
si
ne
mo
c
sp
oj

ov
al
a
se
za
po
mn
ět
li
vo
st
í
ne
bo
zt
rá
to
u
ci
tu
v
ko
nč
et
in
ác
h.
Ja
k
po
st
ih
uj
e
hl
as
?

V řeči se velmi často projevuje pozměněná výslovnost hlásek, přerývané dýchání, únava, něco je tam často jinak. Ovlivňuje to artikulaci, a tím pádem i řečový signál. A protože jsme schopní měřit akustický signál pomocí mikrofону, můžeme tak nepřímo měřit i roztroušenou sklerózu.

Takže ve výzkumu vycházíte z toho, že u většiny případů s roztroušenou sklerózou se nemoc projeví v hlase.

Ano. Když ale říkáme ve většině případů, tak záleží, koho se ptáte. Existuje spousta poslechových i akustických studií, jak často v hlase roztroušená skleróza je. Střízlivý odhad je, že jde o něco málo přes polovinu případů pacientů, ale já si optimisticky myslím, že je to víc.

Co všechno dokáže hlas člověka prozradit o jeho zdravotním stavu?

Kromě zdraví dnes už dost spolehlivě rozpoznáme mužský a ženský hlas, dokážeme identifikovat mluvčího podle referenční nahrávky z různých bezpečnostních systémů, případně také věk. Pokud jde o zdraví, poznáme intoxikaci, emoční afekt nebo další neurologické choroby, například Parkinsonovu chorobu, s jejímž výzkumem velmi pohnuli kolegové z ČVUT – docent Jan Ruz, doktorka Tereza Tykalová a další.

Když se v hlase projeví roztroušená skleróza, není už pozdě? Může rozpoznání příznaků lidem pomoci včas?

Teď narážíte na hlavního strašáka mé diplomové práce. Lidé, jejichž hlas jsme prostřednictvím nahrávek zkoumali, mají RS ve středně rozvinuté fázi. A právě to vznáší otázku, jestli jde tento výzkum dělat dostatečně brzy na to, aby byl užitečný pro diagnostiku.

Ta dnes bohužel trvá pořád příliš dlouho a v momentě potvrzení diagnózy bývá RS často už ve středně rozvinuté fázi. Dalším problémem je finanční náročnost vyšetření, zejména v zemích jako USA jde o nákladnou záležitost a lidé se jí spíše

vyhýbají, pokud na tom nejsou už vyloženě zle. Náš výzkum může pomoci právě v tomto ohledu, protože potenciálně dotlačí lidi, kteří si nejsou úplně jistí, aby se nechali pořádně vyšetřit.

Můžeme se také bavit o takzvané treatment response, tedy reakci na nově zvolenou léčbu, která se ovšem u každého nemocného liší, a proto trvá, než se potvrdí účinnost léků. Sledování roztroušené sklerózy v hlase před a po užívání léků představuje indikátor, jehož prostřednictvím se dá ověřovat, jestli konkrétní léčba nemocnému pomáhá a příznaky v jeho hlase mizí či ne. Případně můžeme díky této metodě a analýze nahrávek sledovat postup nemoci.

Takže jde o víceúčelovou metodu použitelnou pro různé fáze onemocnění.

Ano, ale vždy by se musela přizpůsobit tomu, co přesně s ní chcete sledovat. Výzkum z mé diplomky nám zatím ukázal, jak dobře by to mohlo fungovat a které akustické parametry jsou při rozboru nahrávek a pro celou metodu významné a které nám toho o nemoci prozradí víc.

Zmiňujete takzvané akustické parametry – co přesně v hlase pacientů s roztroušenou sklerózou hledáte?

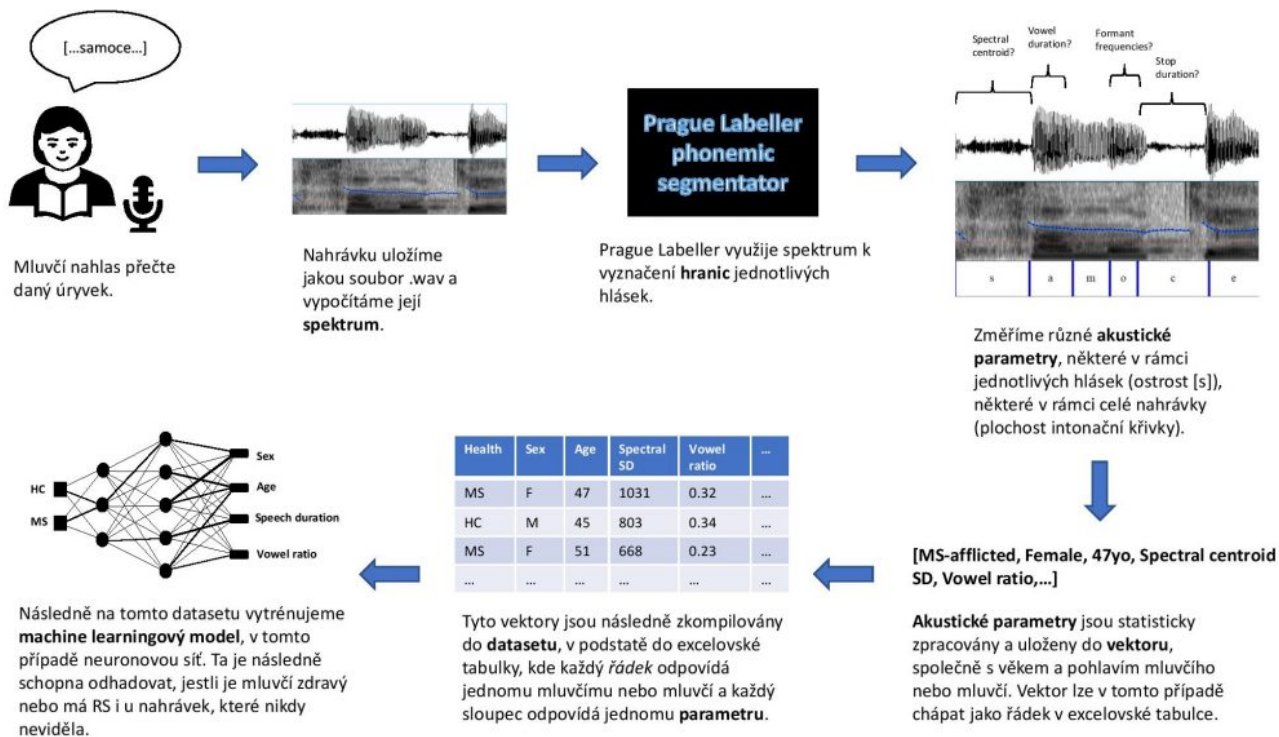
Pacienti, s nimiž jsem pracoval, četli nahlas Čapkův text *Měl jsem psa a kočku* a já pak v nahrávkách měřil vybrané aspekty, které nám o stavu člověka dokážou hodně říct. Celkově jsme měřili třináct akustických parametrů. Například čas, který trvá pacientovi úryvek přečíst – což je samozřejmě skoro až hloupě jednoduché změřit –, je jedním ze statisticky významných parametrů. Sklerotikům trvá čtení déle než zdravému člověku, protože mohou mít potíže se zrakem nebo se dříve unaví, což se projeví právě i v řeči. Druhým parametrem je například intonační křivka, kterou mají sklerotici z důvodu horší koordinace hlasivek plošší, a tím pádem drží intonaci v jednodušší hladině. Dalším aspektem je ostrost výslovnosti souhlásky [s], respektive její směrodatná odchylka. Dalším

symptomem jsou svalové záškuby, takzvaná spasticita, kvůli níž drží při vyslovování závěr některých hlásek, konkrétně [k], [t], [p], déle.

Jde o nepatrné rozdíly, které v běžné řeči často vůbec neslyšíme, protože mozek je filtruje jako bezvýznamný šum. Když je ale naměříme, zjistíme, že mohou být významné a že v kombinaci s dalšími parametry mohou indikovat, že něco je špatně.

Co následuje po pořízení nahrávek?

Potom, co si přečteme nahrávku, ji vložíme do speciálních programů, které z ní vytáhnou informace o intonační křivce, akustickém parametru intenzity, tedy hlasitosti, a jednotlivých realizací fonémů čili o trvání a kvalitě hlásek. Informace zanalyzujeme a zadáme je do strojového učení. Počítači řekneme „postav takový a takový machinelearningový model“ – například neuronovou síť s určitým počtem vrstev –, a model vytrénujeme. Na části dat, které jsme před tréninkem nechali stranou a algoritmus je tedy nikdy neviděl, pak ověřujeme, jestli se je model naučil rozeznávat, a zjišťujeme, jak přesný je. Nám vycházela přibližně 70% úspěšnost, takzvaná balanced accuracy, což je pěkný výsledek, ale prozatím je pouze orientační, protože jsme neměli k dispozici velké množství dat.



Jak se to dělá? Takhle: flowchart výzkumu Emila Svobody.

Váš výzkum je velmi interdisciplinární a přesahuje i moje představy o tom, čemu se věnují studenti fonetiky. Prováděl vás celým procesem někdo?

Na Fonetickém ústavu FF UK působí doktor Tomáš Bořil, vystudovaný řečový technolog, který všem těmto věcem rozumí – od akustiky až po programování. Můj výzkum je poměrně vzácně technický, ale i tak se od vás na fonetice očekává, že se naučíte programovat. Není to kdovíjak náročné. Studenti se učí hlavně metodiku, jak trénovat a vyhodnocovat algoritmy strojového učení. Samozřejmě když nechcete, nemusíte se tím do hloubky zabývat, ale rozhodně k tomu máte příležitost.

I když váš výzkum zatím nejde využít v praxi, má do budoucna velký potenciál. Co s ním plánujete dál?

Společně s docentem Ruszem a doktorkou Tykalovou z ČVUT a doktorem Bořilem z FF UK pokračujeme v další podobné studii. Navazujeme na obsah mé diplomky a rozvádíme ho dál. Vyrovnali jsme všechny metodologické problémy, které mohly výzkum zkreslovat. Zásadní změnou je to, že ve výzkumu teď máme ke každému nemocnému člověku jednoho zdravého člověka, a to ve

zhruba stejném věku a stejného pohlaví. Došli jsme k podobným výsledkům, které hodláme publikovat jako odborný článek. Ale zatím na tom stále ještě pracujeme.

Do budoucna chci nicméně pokračovat ve zkoumání hlasu a toho, co všechno se z něj dá zjistit. Existují totiž pokročilejší metody strojového učení zvané deep learning: tyto neuronové sítě dokáží kompletně samy zanalyzovat nahrávky mluvčích, které jsme dosud procházeli ručně. Jenže aby to mohlo fungovat, potřebujete mnohem víc dat, na nichž se sítě vytrénují. Na tomto principu fungují například automatické titulky na YouTube. Něco podobného chceme udělat, ale zatím jsme vždy měli k dispozici „jen“ pár stovek nahrávek a doteď vlastně netušíme, jestli to bude fungovat i na detekci roztroušené sklerózy. Na to se chci soustředit.

Máte představu, kdy by se program mohl dostat třeba do ordinací lékařů?

Otázka implementace je různá, protože existuje víc možností. Vy zmiňujete využití u doktora, čehož zatím nejsem velkým zastáncem. Přinesla by totiž ordinacím administrativní zátěž spojenou s vyřizováním správných licencí programu. Nicméně teoreticky by mohl sloužit k jakési triáži, která by mohla rozhodovat o tom, koho poslat na magnetickou rezonanci, případně s jakou prioritou. Další možností je vznik aplikace v mobilním telefonu...

Další otázkou je, jakou fyzickou podobu by ta metoda měla vůbec mít – měla by to být nějaká krabička nebo program v počítači? Ale co když doktor nemá správný počítač nebo mikrofon? Váže se k tomu celá řada zatím nevyřešených problémů, kterými se v téhle fázi ještě nezabýváme. Zatím musíme zjistit, jestli to vůbec půjde a rozvinout základní výzkum. Řekl bych, že to není dál za hranicí dekády, ale rozhodně zatím neřeknu, jestli to bude trvat dva roky nebo třeba deset let.

V budoucnu se může ukázat i to, že se dysartrie, tedy onemocnění řeči, potvrdí třeba jen u 10 procent sklerotiků v nejranějších fázích. V momentě, kdy se tohle stane, projekt skončí. Nechci nikoho zklamat, naděje je potřeba, ale zároveň není dobré se na ni upínat. Nic nefunguje rovnou z laboratoře, kde se zrovna nacházíme. Takže uvidíme.

Autorka: Tereza Šindelářová

Článek vznikl pro fakultní magazín Small Talk.