

# Úkol k přijímacímu řízení – Phonexia

1. března 2024

Jan Zdeněk

## 1 Úkol 2 – replikace výsledků pro baseline

K řešení druhého úkolu byla vybrána baseline implementace ASVspoof 2021 Baseline-LFCC-GMM.<sup>1</sup> Pro extrakci příznaků (feature extraction) je zde využíváno metody LFCC (linear frequency cepstral coefficients). Model je trénován pomocí GMM (gaussian mixture model). Následná evaluace (výpočet ERR a tDCF) proběhla pomocí skriptů, které se nacházejí ve stejném repozitáři jako implementace baseline řešení. Ve článku ASVspoof 2021: Towards Spoofed and Deepfake Speech Detection in the Wild<sup>2</sup> je tato implementace označována jako B02.

—	C1	C2	C3	C4	C5	C6	C7	Pooled
A07	0.4935	0.6572	0.8429	0.5444	0.7036	0.6275	0.4523	0.6505
A08	0.0699	0.3067	0.4586	0.0926	0.3330	0.3651	0.0686	0.2689
A09	0.2410	0.3825	0.5871	0.2706	0.3477	0.4131	0.2694	0.4315
A10	0.5779	0.6878	0.8727	0.6538	0.7608	0.6392	0.5082	0.7067
A11	0.1973	0.5444	0.9060	0.2841	0.6061	0.5805	0.2732	0.5149
A12	0.1299	0.9348	0.9625	0.3756	0.5967	0.8043	0.1515	0.6072
A13	0.1224	0.4203	0.6475	0.1769	0.4511	0.4418	0.1495	0.3690
A14	0.1988	0.4152	0.7176	0.2703	0.4586	0.4412	0.1928	0.4094
A15	0.2370	0.5037	0.7530	0.2834	0.5390	0.4629	0.2281	0.4543
A16	0.2002	0.7353	0.8699	0.3872	0.6515	0.6882	0.2453	0.5693
A17	0.9177	0.9738	0.9553	0.9187	0.9835	0.9011	0.8800	0.9243
A18	0.9803	0.9512	0.9901	0.9934	0.9641	0.9297	0.9023	0.9779
A19	0.9018	0.9359	0.9003	0.9249	0.9530	0.9352	0.9314	0.9472
Pooled	0.3867	0.6394	0.8068	0.4588	0.6404	0.6374	0.3810	<b>0.5944</b>

Tabulka 1: Replikované hodnoty tDCF po evaluaci natrénovaného baseline modelu. Tučně je vyznačena hodnota k porovnání s referenční tabulkou 3.

—	C1	C2	C3	C4	C5	C6	C7	Pooled
A07	18.03	23.34	32.68	20.56	25.91	21.38	16.84	23.02
A08	0.36	6.35	11.80	1.24	7.89	7.34	0.50	5.34
A09	0.06	2.14	11.51	0.15	0.98	2.11	0.05	3.05
A10	22.12	25.46	32.78	24.82	29.03	22.15	19.22	25.61
A11	6.52	20.00	33.42	9.82	20.93	19.17	9.55	17.17
A12	3.34	37.32	38.13	13.06	19.91	28.91	4.02	21.38
A13	2.53	10.59	18.74	4.07	11.46	10.01	3.78	8.72
A14	6.15	12.23	25.91	8.47	13.83	11.49	5.85	12.18
A15	7.49	16.49	26.05	9.24	18.25	12.74	7.35	14.13
A16	7.19	26.93	32.05	13.69	22.95	24.08	8.73	19.44
A17	17.38	25.62	31.41	18.54	26.64	27.93	21.32	24.29
A18	20.47	22.13	39.82	25.39	20.93	24.53	17.77	25.96
A19	32.02	30.88	36.81	29.63	33.28	31.37	34.29	34.50
Pooled	13.74	21.46	29.26	16.05	20.50	20.12	13.62	<b>19.42</b>

Tabulka 2: Replikované hodnoty EER po evaluaci natrénovaného baseline modelu. Tučně je vyznačena hodnota k porovnání s referenční tabulkou 3.

<sup>1</sup>Dostupné z: <https://github.com/asvspoof/asvspoof-challenge-2021/tree/main/LA/Baseline-LFCC-GMM/python>

<sup>2</sup>Dostupné z: <https://arxiv.org/pdf/2210.02437.pdf>

#	ID	tDCF	EER
1	T23	0.2177	1.32
2	T35	0.2480	2.77
3	T19	0.2495	3.13
4	—	0.2500	2.81
5	T36	0.2531	3.10
6	T20	0.2608	3.21
7	T08	0.2672	3.62
8	T16	0.2689	3.63
9	—	0.2725	3.61
10	T04	0.2747	5.58
11	T06	0.2853	5.66
12	—	0.2880	5.01
13	T03	0.2882	4.66
14	—	0.2893	5.70
15	T31	0.3094	5.46
16	T17	0.3279	7.19
17	T07	0.3310	8.23
18	T30	0.3362	8.89
19	B03	0.3445	9.26
20	T02	0.3446	7.79
21	T14	0.3451	8.98
22	T11	0.3666	7.19
23	T34	0.4059	13.45
24	B04	0.4257	9.50
25	T15	0.4890	14.68
26	B01	0.4974	15.62
27	T25	0.5148	13.75
28	T32	0.5270	12.90
29	—	0.5748	18.50
<b>30</b>	<b>B02</b>	<b>0.5758</b>	<b>19.30</b>
31	—	0.5775	14.28
32	T01	0.6204	15.95
33	—	0.6288	15.87
34	T24	0.6320	15.98
35	—	0.6371	16.27
36	T29	0.6741	17.41
37	—	0.6813	17.66
38	T12	0.7228	26.41
39	—	0.7233	19.19
40	—	0.8521	26.14
41	—	1.0000	53.81

Tabulka 3: Referenční hodnoty pro baseline modely (B01 - B04) a modely odevzdané v rámci ASVspoof2021. Tučně jsou vyznačeny referenční hodnoty pro porovnání s replikovanými hodnotami z tabulek 1 a 2.

## 2 Úkol 3 – identifikace anonymizovaného modelu

Pro úkol 3 nebylo nalezeno již natrénovaného modelu a tudíž byla vybrána jiná dostupné implementace a to **ASVspoof2021-ECAPA-TDNN**.<sup>3</sup> Vzhledem k časové náročnosti trénování modelů nebylo možné vše dokončit a tedy nebude ani možné identifikovat o který anonymizovaný model z výše uvedeného článku se jedná. Celkem bylo natrénováno jen 61 ze 100 modelů (100 je výchozí hodnota skriptu `main.py`), přičemž pomocí skriptu `find_best.py` byl vybrán nejúspěšnější model (model 25<sup>4</sup>). Model s indexem 25 byl poté otestován pomocí skriptu `test.py`, kdy výsledné hodnoty činí **min\_tDCF: 0,7494** a **ERR: 20,16**.

## 3 Porovnání a závěr z úkolů 2 a 3

Model z úkolu 3 by se potenciálně zařadil až na 40. místo, což by jej zařadilo výrazně hůře než jen baseline B02, který je sám o sobě nejhorší baseline implementací. Rozdílné pak byly hodnoty min\_tDCF a ERR, které byly spočteny hned po natrénování modelu (min\_tDCF: 0,4079 a ERR: 11,89), pro tyto hodnoty však bylo využito jiného datasetu.

---

<sup>3</sup>Dostupné z: <https://github.com/alexw914/ASVspoof2021-ECAPA-TDNN>

<sup>4</sup>Dostupné z: <https://github.com/hoza2000/phonexia/tree/main>