ICT응용
기계학습을 이용한 새소리 분류

2011270314

컴퓨터정보학과

서인석

개요

- 목적

새의 소리를 기계학습을 이용하여 학습하여, 새의 종을 분류한다.

- 실험 환경

MATLAB

- 실험 과정

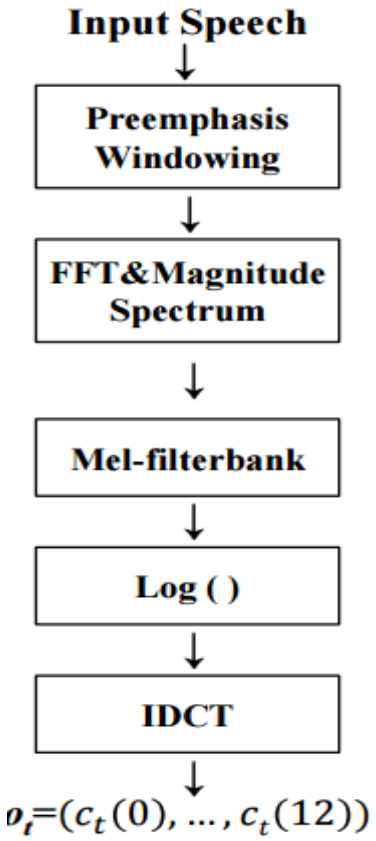
1. 새의 소리를 기반으로 MFCC특징 추출
2. 추출한 MFCC를 전처리 한다
3. 기계학습 알고리즘 중 K-NN(K-Nearest Neighbor) 알고리즘을 이용하여 분류한다
4. 결과 분석

MFCC(Mel-frequency cepstral coefficient)

- 입력된 소리 전체를 대상으로 하는 것이 아니라, 일정 구간 식 나누어, 이 구간에 대한 스펙트럼을 분석하여 특징을 추출하는 기법이다.

- 추출과정

1. 주어진 데이터를 프레이밍하기.
2. DFT를 이용하여 신호를 구한다.
3. 각 신호의 세기를 mel scale에 집어넣는다
4. 각 mel scale 세기의 log를 구한다
5. Mel log파워(log filterbank energe)
   의 IDCT를 구한다



K-NN(K-Nearest Neighbor)

- k-NN 알고리즘은 지도 학습의 한 종류이다.
- 레이블이 있는 데이터를 사용하여 분류 작업을 하는 알고리즘이다.
- 알고리즘의 이름에서 볼 수 있듯이 데이터로부터 거리가 가까운 k개의 다른 데이터의 레이블을 참조하여 분류하는 알고리즘이다.
- 주로 거리를 측정할 때 유클리디안 계산법을 사용하여 거리를 측정한다.

$$\text{Euclidian distances} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Train dataset

Acorn_Woodpecker_00001	Acorn Woodpecker (Mela...	Eric Cannizzaro	xeno-canto
Acorn_Woodpecker_00002	Acorn Woodpecker (Mela...	Eric Cannizzaro	xeno-canto
Acorn_Woodpecker_00003	Acorn Woodpecker (Mela...	Juan Carlos P챕re...	xeno-canto
Acorn_Woodpecker_00004	Acorn Woodpecker (Mela...	David Vander Plu...	xeno-canto
Allen_Hummingbird_00001	Allen's Hummingbird (Se...	Thomas G. Graves	xeno-canto
Allen_Hummingbird_00002	Allen's Hummingbird (Se...	Richard E. Webster	xeno-canto
Allen_Hummingbird_00003	Allen's Hummingbird (Se...	Richard E. Webster	xeno-canto
Allen_Hummingbird_00004			
American_Avocet_00001			
American_Avocet_00002	American Avocet (Recurv...	Tim Marquardt	xeno-canto
American_Avocet_00003	American Avocet (Recurv...	Tim Marquardt	xeno-canto
American_Avocet_00004	American Avocet (Recurv...	Richard E. Webster	xeno-canto
American_Black_Oystercatcher_00001	Black Oystercatcher (Hae...	Ian Cruickshank	xeno-canto
American_Black_Oystercatcher_00002	Black Oystercatcher (Hae...	Paul Marvin	xeno-canto
American_Black_Oystercatcher_00003	Black Oystercatcher (Hae...	Paul Marvin	xeno-canto
American_Black_Oystercatcher_00004	Black Oystercatcher (Hae...	Ian Cruickshank	xeno-canto
American_Rock_Wren_00001			
American_Rock_Wren_00002	Rock Wren (Salpinctes o...	Elisa Yang	xeno-canto
American_Rock_Wren_00003	Rock Wren (Salpinctes o...	Elisa Yang	xeno-canto
American_Rock_Wren_00004	Rock Wren (Salpinctes o...	Patrick Turgeon	xeno-canto
Baird_Sandpiper_00001	Baird's Sandpiper (Calidri...	Harry Lehto	xeno-canto
Baird_Sandpiper_00002	Baird's Sandpiper (Calidri...	Andrew Spencer	xeno-canto
Baird_Sandpiper_00003	Baird's Sandpiper (Calidri...	Andrew Spencer	xeno-canto

Test dataset

Acorn_Woodpecker_00005	Acorn Woodpecker (Mela...	Mario Trejo	xeno-canto
Allen_Hummingbird_00005	Allen's Hummingbird (Se...	Eric DeFonso	xeno-canto
American_Avocet_00005	American Avocet (Recurv...	Micah Riegner	xeno-canto
American_Black_Oystercatcher_00005	Black Oystercatcher (Hae...	Ian Cruickshank	xeno-canto
American_Rock_Wren_00005	Rock Wren (Salpinctes o...	Patrick Turgeon	xeno-canto
Baird_Sandpiper_00005	Baird's Sandpiper (Calidri...	Andrew Spencer	xeno-canto
Band-tailed_Pigeon_00005	Band-tailed Pigeon (Pata...	Peter Boesman	xeno-canto
Black_Phoebe_00005	Black Phoebe (Sayornis n...	Juan Carlos P챕re...	xeno-canto
Black_Skimmer_00005	Black Skimmer (Rynchop...	Aidan Place	xeno-canto
Black_Turnstone_00005	Black Turnstone (Arenaria...	Elias Aristides Elias	xeno-canto
Burrowing_Owl_00005	Burrowing Owl (Athene c...	Jo찾o Ant척nio d...	xeno-canto
Royal_Tern_00005	Royal Tern (Thalasseus m...	Paul Marvin	xeno-canto
Surfbird_00005	Surfbird (Aphriza virgata)	Mike Nelson	xeno-canto

MFCC

MFCC_train | MFCC_train(1).MFCCs | MFCC_test | mfcc_Train

1x52 struct 3개 필드 포함

필드	MFCCs	FBEs	frames
1	13x757 double	20x757 double	1103x757 double
2	13x5305 double	20x5305 double	1103x5305 double
3	13x3244 double	20x3244 double	1103x3244 double
4	13x2740 double	20x2740 double	551x2740 double
5	13x3105 double	20x3105 double	1103x3105 double
6	13x2614 double	20x2614 double	1103x2614 double
7	13x1938 double	20x1938 double	1103x1938 double
8	13x1072 double	20x1072 double	1103x1072 double
9	13x1098 double	20x1098 double	1103x1098 double
10	13x3048 double	20x3048 double	1103x3048 double
11	13x3040 double	20x3040 double	1103x3040 double
12	13x3163 double	20x3163 double	1103x3163 double
13	13x597 double	20x597 double	1200x597 double
14	13x6215 double	20x6215 double	1200x6215 double
15	13x4113 double	20x4113 double	1200x4113 double
16	13x1550 double	20x1550 double	1200x1550 double
17	13x4589 double	20x4589 double	1103x4589 double
18	13x1465 double	20x1465 double	1103x1465 double
19	13x1661 double	20x1661 double	1103x1661 double
20	13x5759 double	20x5759 double	1103x5759 double
21	13x2013 double	20x2013 double	1103x2013 double
22	13x2164 double	20x2164 double	1200x2164 double

- Train 데이터 셋을 기반으로 MFCC특징을 추출

MFCC

50.0524	50.5748	50.6095	50.8970	50.8032	51.0252	50.6525	52.1894	51.0697	50.8952	51.1273	51.0719	49.0523	51.6005
-1.3783	-1.9266	-4.0167	-4.3001	-4.5507	-2.4320	-1.3690	-2.3460	-3.8964	-5.6343	-3.3083	-2.0793	-5.9597	-2.3962
-0.5758	1.7391	-1.2315	1.7337	4.7770	4.0540	4.2938	2.7542	2.4644	2.9196	3.3344	3.8553	3.9065	5.0899
2.7805	3.2256	3.2156	6.5301	4.6350	5.1773	2.7373	2.5404	1.3481	2.5670	4.3615	3.2134	3.6648	0.1864
6.2463	6.2847	4.3447	-0.0358	7.0334	4.2962	1.3789	4.2435	4.9537	1.9574	5.7416	6.5437	4.8052	-0.0199
3.6301	2.4057	0.8918	0.7541	4.5565	-1.6799	3.8243	-0.2422	1.3138	-1.3622	-0.7450	4.2629	5.9528	0.9300
1.9674	1.0184	-1.0559	-2.4657	-0.0175	-3.5023	8.5000	-0.9295	-1.4098	0.1586	1.9439	0.7224	3.2669	0.7465
-0.4655	3.3732	1.8542	-5.1753	-3.9109	-0.0379	-2.2053	1.0675	0.5408	-7.0062	-2.7606	0.9888	0.1582	-0.1496
0.3925	4.5618	-2.6653	-4.1886	0.0383	-0.7377	-2.1833	-2.5265	1.5028	-5.4235	-0.6131	2.6702	3.5956	-1.6905
-10.2438	-2.1341	-3.0367	-3.7731	3.5903	-0.4706	-1.2101	0.2592	-0.8600	2.2883	-5.6176	-10.2471	-4.2773	-2.6897
2.1280	1.2960	3.7403	5.0455	1.5345	-2.2471	-2.6699	4.3806	-0.0182	-1.8758	-5.1244	-6.7264	-13.2388	-5.5238
-2.9229	-0.2663	4.4730	5.0762	4.8970	-2.0082	1.0980	1.2590	6.5426	2.8554	0.5065	7.4345	6.0327	-1.3618
-3.1628	-2.0261	-0.1354	1.6529	-2.7379	1.1423	1.2462	-0.1932	-1.7117	-1.9602	0.3309	-2.3389	0.5790	-0.5073

Acorn_Woodpecker_00001 파일에서 추출된 MFCC (Cepstral 상수의 수는 13)

11.1322	19.0754	23.5692	51.8594	54.0435	54.6797	53.1201	51.8563	52.2784	51.9436	52.3953	52.8868	53.5231	54.1309
-2.6081	-1.8258	-1.1471	-0.2984	1.7252	5.1406	3.2421	2.9137	1.4601	-0.0013	3.7025	3.3646	4.3397	2.8867
3.3526	1.1231	1.9722	1.0892	3.8146	5.8204	8.1887	7.6649	4.1593	3.2775	4.3881	4.0392	5.9741	3.1499
-3.1863	-4.3046	-2.7885	-1.6501	-0.6753	-1.9980	-5.2579	-1.4779	-7.3935	-7.3855	-4.6968	-4.3445	-4.6910	-5.1364
-4.4645	-6.0273	1.5668	-2.9270	-4.6648	-6.5578	-6.5088	-3.9360	-5.5413	-4.0551	-5.4817	-7.4524	-2.5029	-2.8192
-2.9083	-4.0126	-0.8154	1.2452	-1.7834	-10.5321	-2.4209	-4.3216	-0.0075	1.4756	-2.5556	-2.3073	0.1224	-2.9591
5.3808	-1.5495	-0.5561	-1.4560	-3.8103	-2.8153	-1.2090	0.1283	2.7778	2.0703	2.0320	-1.6878	-0.3020	0.9534
1.8991	4.3636	-6.0536	-1.5684	0.3238	5.8509	2.2685	-2.6851	1.9097	3.2633	1.0518	2.0279	1.4424	-4.3483
-1.3325	0.7286	-1.1813	0.1784	-3.1148	-4.5807	-4.2566	-4.8118	0.0215	0.8180	0.8803	-6.7753	1.6996	-2.7422
-1.3144	2.0997	6.2196	-2.2412	-3.6715	-6.3259	-1.3750	-4.1278	0.2426	-1.9517	-1.0986	-1.9079	1.0581	-0.5095
-1.1800	6.4664	3.4928	-0.3930	-1.6936	-5.8853	-1.6772	4.9708	5.0862	-0.4675	1.2315	3.8901	8.3704	0.9532
-2.6943	-4.6195	0.9725	-6.1203	-4.4836	-5.6298	-1.0116	4.1656	5.6302	2.3262	4.7581	1.9744	2.3482	4.6784
0.6426	0.8916	-8.2450	1.9975	0.3359	1.2412	3.1076	5.0914	-1.8596	-0.7126	4.4767	-1.2421	-2.9352	-1.4520

American_Rock_Wren_00002 파일에서 추출된 MFCC (Cepstral 상수의 수는 13)

MFCC (Z_SCORE STANDARDIZATION)

3.1973	3.2697	3.2591	3.2072	3.2119	3.2649	3.2417	3.2892	3.2651	3.2362	3.2335	3.1310	3.0746	3.282
-0.3523	-0.5169	-0.5923	-0.5661	-0.7069	-0.4501	-0.4466	-0.4963	-0.6101	-0.5919	-0.4843	-0.4475	-0.7155	-0.394
-0.2970	-0.2525	-0.3959	-0.1536	-0.0465	6.5926e-04	-0.0451	-0.1423	-0.1616	-0.0126	-0.0306	-0.0479	-0.0358	0.115
-0.0653	-0.1453	-0.0824	0.1743	-0.0566	0.0787	-0.1555	-0.1571	-0.2403	-0.0365	0.0395	-0.0911	-0.0524	-0.218
0.1739	0.0753	-0.0027	-0.2746	0.1132	0.0175	-0.2518	-0.0389	0.0139	-0.0778	0.1338	0.1331	0.0261	-0.232
-0.0067	-0.2045	-0.2462	-0.2206	-0.0622	-0.3978	-0.0784	-0.3503	-0.2427	-0.3026	-0.3092	-0.0205	0.1052	-0.168
-0.1214	-0.3045	-0.3835	-0.4407	-0.3860	-0.5245	0.2531	-0.3980	-0.4348	-0.1996	-0.1256	-0.2588	-0.0799	-0.180
-0.2893	-0.1347	-0.1783	-0.6259	-0.6616	-0.2837	-0.5059	-0.2594	-0.2972	-0.6848	-0.4469	-0.2409	-0.2940	-0.241
-0.2301	-0.0489	-0.4970	-0.5585	-0.3820	-0.3323	-0.5043	-0.5089	-0.2294	-0.5776	-0.3002	-0.1277	-0.0572	-0.346
-0.9642	-0.5319	-0.5232	-0.5301	-0.1306	-0.3138	-0.4353	-0.3155	-0.3960	-0.0554	-0.6420	-0.9974	-0.5996	-0.414
-0.1103	-0.2845	-0.0454	0.0728	-0.2761	-0.4372	-0.5388	-0.0294	-0.3367	-0.3374	-0.6083	-0.7604	-1.2170	-0.607
-0.4590	-0.3972	0.0063	0.0749	-0.0381	-0.4206	-0.2717	-0.2461	0.1259	-0.0170	-0.2238	0.1931	0.1107	-0.324
-0.4755	-0.5241	-0.3186	-0.1591	-0.5786	-0.2017	-0.2612	-0.3469	-0.4560	-0.3431	-0.2358	-0.4650	-0.2650	-0.266

Acorn_Woodpecker_00001 파일에서 추출된 MFCC (Cepstral 상수의 수는 13)

3.2338	3.1714	3.2461	3.2475	3.2697	3.2707	3.2495	3.2167	3.2546	3.2467	3.2569	3.2879	3.2774	3.21
-0.2733	-0.2937	-0.2989	-0.1846	-0.1667	-0.2168	-0.2659	-0.3432	-0.2820	-0.3513	-0.2879	-0.3198	-0.2479	-0.28
-0.3519	-0.4153	-0.3731	-0.4407	-0.2239	-0.2024	-0.1672	-0.3419	-0.2947	-0.1024	-0.3564	-0.3084	-0.1916	-0.41
-0.0249	-0.0510	-0.0135	-0.1533	-0.2858	-0.2912	0.0047	-0.1242	0.0114	-0.0661	-0.2749	-0.2949	-0.1933	-0.10
-0.1470	0.2685	0.2092	0.2207	-0.0575	-0.1139	-0.2704	-0.0785	0.0512	-0.0058	0.2342	-0.2035	0.1151	0.15
-0.7590	-0.4722	-0.5379	-0.5512	-0.7069	-0.7141	-0.7652	-0.5453	-0.4882	-0.5818	-0.6618	-0.5907	-0.5318	-0.71
-0.4545	-0.0659	-0.4448	-0.1887	-0.2105	-0.2958	-0.1089	-0.4295	-0.1687	-0.2281	-0.1158	-0.1794	-0.2846	-0.25
0.0839	0.1470	-0.4194	-0.0850	-0.2438	0.0207	-0.2398	-0.3558	-0.4654	0.0409	-0.1691	-0.4100	-0.1778	-0.18
-0.0857	-0.2094	-0.2593	-0.5141	-0.0566	-0.2067	-0.6169	-0.5128	-0.7211	-0.4376	-0.4669	-0.3330	-0.6040	-0.34
-0.5787	-0.5477	-0.3533	-0.5035	-0.5448	-0.5466	-0.3329	-0.6896	-0.3591	-0.7215	-0.1795	-0.0751	-0.3450	-0.55
-0.2326	-0.9012	-0.0415	-0.2823	-0.4271	-0.3234	-0.0260	-0.0858	-0.1077	-0.2470	-0.2974	-0.0195	-0.2115	-0.59
-0.0450	-0.2661	-0.1580	-0.4400	-0.1845	-0.1791	-0.1427	0.1582	-0.2481	-0.3293	-0.3144	-0.1274	-0.2635	0.00
-0.3651	-0.3645	-0.5556	-0.1248	-0.1617	-0.2014	-0.3186	0.1317	-0.1822	-0.2167	-0.3669	-0.4262	-0.3415	0.08

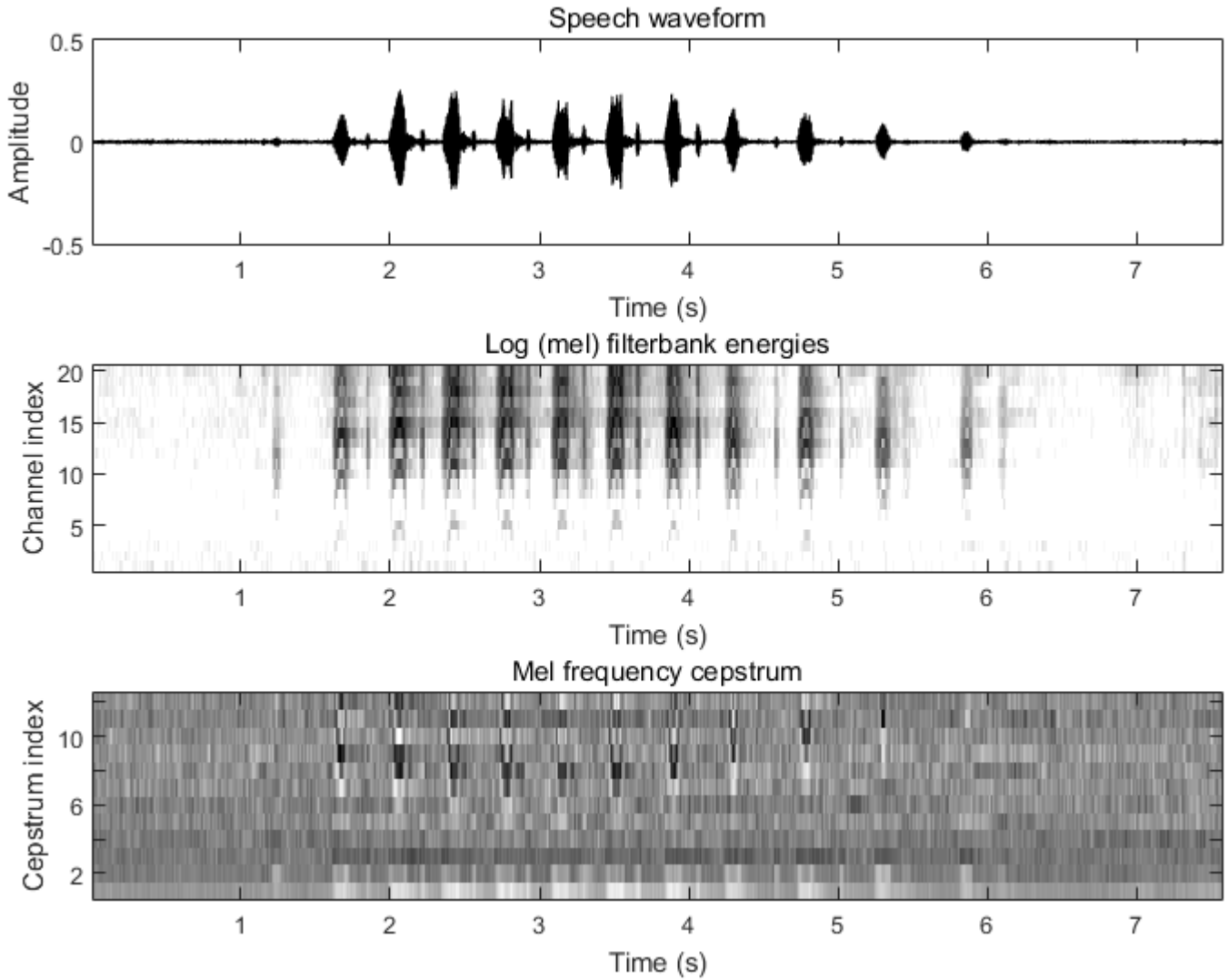
American_Rock_Wren_00002 파일에서 추출된 MFCC (Cepstral 상수의 수는 13)

MFCC

-6.4051e-17	3.4161e-17	-7.6862e-17	-1.0675e-17	5.1241e-17	-2.1350e-17	-4.8038e-18	2.1350e-17	-4.2701e-17	-5.1241e-17	4.2701e-18	-7.2592e-17	-5.1241e-17	-1.0675e-17
5.9781e-17	9.6077e-18	-8.5402e-17	-6.6186e-17	-4.4836e-17	5.2309e-17	6.7254e-17	-1.4945e-17	5.1241e-17	1.0141e-17	-1.2383e-16	-5.7646e-17	7.1524e-17	-4.4836e-17
NaN	NaN	-5.1241e-17	5.1241e-17	-1.0675e-17	4.2701e-18	-2.5621e-17	-8.5402e-17	-9.8212e-17	2.5621e-17	-8.5402e-18	3.5762e-17	8.5402e-18	-1.4945e-17
6.4051e-18	-4.4302e-17	-4.0566e-17	4.2701e-18	-1.1956e-16	-1.0248e-16	-5.7646e-17	4.2701e-18	-1.4945e-17	-1.7080e-17	-5.9781e-17	-1.6226e-16	5.1241e-17	1.6013e-17
2.1350e-17	2.9891e-17	3.2026e-17	-1.0675e-17	2.4153e-17	-8.5402e-18	1.2810e-17	-2.9891e-17	-4.2701e-18	-4.3768e-17	1.2810e-17	-8.5402e-18	6.5653e-17	1.1743e-17
-5.5511e-17	5.9781e-17	-8.8604e-17	2.1350e-18	-8.1132e-17	-2.9891e-17	-2.5621e-17	8.5402e-18	-4.1633e-17	-4.0566e-17	-1.9215e-17	6.4051e-17	8.5402e-18	-3.2026e-18
-2.9891e-17	-3.2026e-18	3.4161e-17	6.4051e-18	-1.8148e-17	-4.2701e-18	1.7080e-17	-1.7080e-17	-1.7080e-17	-3.4161e-17	-7.2592e-17	1.4812e-17	0	1.7080e-17
NaN	7.6862e-17	8.5402e-18	-2.9891e-17	2.1350e-18	4.2701e-18	6.8321e-17	-7.0456e-17	2.1350e-17	-1.2810e-17	1.3878e-17	-4.2701e-18	-3.3360e-17	-3.1492e-17
NaN	NaN	1.7080e-17	3.4161e-17	-8.5402e-17	3.4161e-17	-6.4051e-18	-2.1350e-17	3.2026e-18	2.1350e-18	-1.4945e-17	-1.8148e-17	-1.0141e-17	-4.2701e-18
-5.5511e-17	-1.4945e-17	2.5621e-17	-3.8431e-17	8.3267e-17	4.9106e-17	6.0315e-17	2.1350e-17	-8.5402e-18	-2.6688e-17	-9.3942e-17	-5.5511e-17	5.5511e-17	2.9891e-17
2.5621e-17	1.2810e-17	8.5402e-18	8.5402e-18	5.1241e-17	-2.7756e-17	1.2810e-17	-8.5402e-18	4.6971e-17	6.4051e-17	-2.9891e-17	2.6688e-17	5.1241e-17	1.2810e-16
NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-7.6862e-17	-1.7080e-17
8.5402e-18	0	-3.2026e-17	-2.7756e-17	-9.6077e-18	-1.0248e-16	-3.2026e-18	-1.2810e-17	2.9891e-17	1.4945e-17	-2.3485e-17	4.9106e-17	-5.9781e-17	-1.1102e-16

전처리를 통해 13xN(각각 파일의 길이)의 행렬을 1xN으로 변경 후 변수에 저장

PLOT



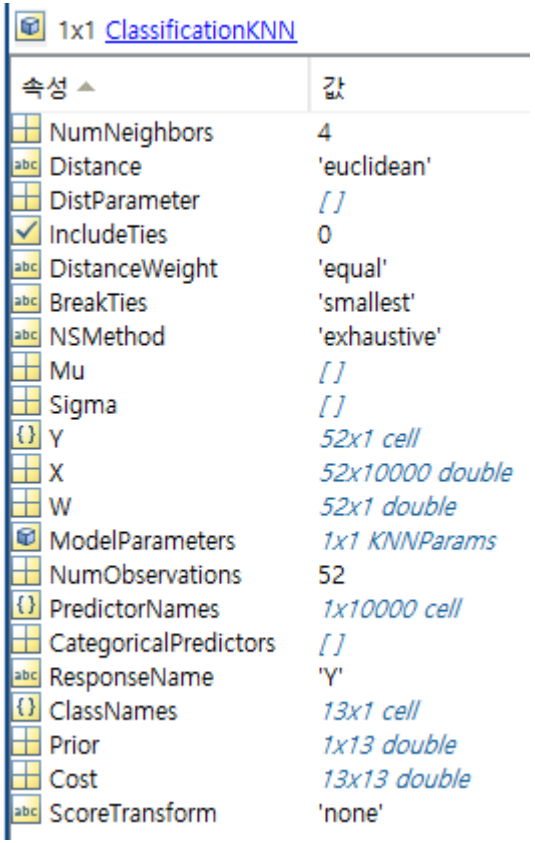
Acorn_Woodpecker_00001.mp3

KNN 모델 구축

```
[num,train_Species] = xlsread('train_species.xlsx');
[num,test_Species] = xlsread('test_species.xlsx');

mdl = fitcknn(mfcc_Train,train_Species,'NumNeighbors',4);

disp('Model Construct');
```



TEST 데이터 셋 입력

Bird_Classification_getdata.m | Bird_Classification_test.m | Bird_Classification_plot.m

```
[BirdClass,score,cost] = predict(mdl,mfcc_Test);

count = 0;
[row,col] = size(BirdClass);

for i = 1:row
    bird = BirdClass{i,1};
    species = train_Species{i,1};
    if(size(bird) == size(species))
        count = count + 1;
    end
end

Rate = (count/ row) * 100
```

```
Rate =

    23.0769
```

생각보다 낮은 적중률을 보인다!

성능향상

- Min_Max nomalize 사용

```
Rate =

    7.6923
```

오히려 성능이 감소

- 데이터의 길이 변경

```
Rate =

   23.0769
```

데이터 길이 10000에서 7000으로 변경했지만 성능은 동일

- Cepstral 상수 크기 변경

```
Rate =

   30.7692
```

성능이 조금이나마 향상

생각보다 낮은 적중률을 보인다. 이유는?

- 부족한 Train 데이터 셋

  기계학습이란 데이터를 이용하여 학습을 하는 것, 즉 데이터가 많을수록 높은 학습률을 보인다.

- 부적절한 전처리 과정

  KNN알고리즘에 적용하기 위해 데이터의 길이를 임의로 선정(7000)하여 길이만큼 자르거나, zero padding을 사용했다. 이 부분에서 데이터의 변형이 일어난다. 구한 또는 13xN의 mfcc 데이터를 1xN의 크기로 때, 산술평균(mean)이 아닌, 다른 수학적인 합당한 방법이 존재 할 수도 있다.

- 부적절한 알고리즘 선정

  새소리의 분류라는 문제에 대한 해결방법으로 KNN알고리즘을 선정했지만, 보다 나은 기계학습 알고리즘이 존재 할 수도 있다.

Q & A