

# Music Dataset classification



AutoML 연구실  
석사 과정 조성현  
21.05.25

# Dataset statistics of single label - 900 features

| Dataset                 | Patterns | Features |         | # of Classes | Average Patterns per Class $\pm$ Standard Deviation | Domain           | 길이 구분           |
|-------------------------|----------|----------|---------|--------------|-----------------------------------------------------|------------------|-----------------|
|                         |          | Size     | Type    |              |                                                     |                  |                 |
| GTZAN                   | 1000     | 888      | Numeric | 10           | 100.000 $\pm$ 0.000                                 | Genre            | 30s             |
| HOMBURG                 | 1886     | 888      | Numeric | 9            | 209.556 $\pm$ 134.866                               | Genre            | 10s             |
| FMA_SMALL               | 7996     | 888      | Numeric | 8            | 999.500 $\pm$ 0.500                                 | Genre            | 30s             |
| MER500                  | 493      | 888      | Numeric | 5            | 98.600 $\pm$ 2.332                                  | Emotion          | 30s $\pm$ 30s   |
| emoMusic                | 1000     | 888      | Numeric | 8            | 125.000 $\pm$ 0.000                                 | Emotion          | 45s             |
| Soundtracks             | 470      | 884      | Numeric | 9            | 52.222 $\pm$ 19.876                                 | Emotion          | 24s $\pm$ 14s   |
| Emotifymusic            | 400      | 888      | Numeric | 4            | 100.000 $\pm$ 0.000                                 | Genre            | 43 $\pm$ 18s    |
| Ballroom                | 698      | 888      | Numeric | 8            | 87.250 $\pm$ 17.555                                 | Genre            | 30s             |
| SeyerlehnerUnique       | 3115     | 888      | Numeric | 10           | 311.500 $\pm$ 248.349                               | Genre            | 30s             |
| MIREX-like_mood         | 903      | 884      | Numeric | 28           | 32.250 $\pm$ 4.778                                  | Mood             | 30s             |
| Medley-solos            | 21571    | 873      | Numeric | 8            | 2537.875 $\pm$ 2016.080                             | Instrument       | 3s              |
| ISMIR04                 | 727      | -        | Numeric | 6            | 121.167 $\pm$ 95.602                                | Genre            | 30s             |
| good-sounds             | 8399     | 865      | Numeric | 5            | 549.400 $\pm$ 198.697                               | Instrument       | 9s $\pm$ 9s     |
| MICM                    | 1137     | 888      | Numeric | 7            | 162.429 $\pm$ 119.717                               | Modal System     | 6m $\pm$ 6m     |
| Tropical Genres         | 1500     | 888      | Numeric | 5            | 300.000 $\pm$ 0.000                                 | Genre            | 30s             |
| PCMIR                   | 2410     | 886      | Numeric | 6            | 401.667 $\pm$ 46.241                                | Instrument       | 5s              |
| extendedBallroom        | 4180     | 888      | Numeric | 13           | 321.538 $\pm$ 195.702                               | Genre            | 2m $\pm$ 2m     |
| giantsteps_tempo        | 664      | 888      | Numeric | 23           | 28.870 $\pm$ 31.872                                 | Tempo            | 1m30s $\pm$ 30s |
| giantsteps_key          | 604      | 888      | Numeric | 24           | 25.167 $\pm$ 20.418                                 | Key              | 1m20s $\pm$ 40s |
| GMD                     | 1042     | 888      | Numeric | 18           | 57.889 $\pm$ 70.069                                 | Genre            | 5m $\pm$ 5m     |
| FMA_MEDIUM              | 24984    | 888      | Numeric | 16           | 1561.500 $\pm$ 2069.356                             | Genre            | 30s             |
| IRMAS                   | 6700     | -        | Numeric | 5            | -                                                   | Instrument       | 2s              |
| MagnaTagATune_emotion   | 25860    | 888      | Numeric | 17           | -                                                   | Emotion          | 30s             |
| MagnaTagATune_genre     | 25860    | 888      | Numeric | 58           | -                                                   | Genre            | 30s             |
| MagnaTagATune_top50     | 25860    | 888      | Numeric | 45           | -                                                   | Instrument       | 30s             |
| MagnaTagATune           | 25860    | 888      | Numeric | 9            | -                                                   | Tempo            | 30s             |
| CAL500_emotion          | 498      | 1277     | Numeric | 27           | -                                                   | Emotion          | 11m $\pm$ 11m   |
| CAL500_instrument&vocal | 482      | 888      | Numeric | 14           | -                                                   | Instrument&Vocal | 11m $\pm$ 11m   |
| CAL500_genre            | 449      | 888      | Numeric | 12           | -                                                   | Genre            | 11m $\pm$ 11m   |

# Dataset statistics of single label - 400 features

| Dataset                 | Patterns | Features |         | # of Classes | Average Patterns per Class $\pm$ Standard Deviation | Domain           | 길이 구분           |
|-------------------------|----------|----------|---------|--------------|-----------------------------------------------------|------------------|-----------------|
|                         |          | Size     | Type    |              |                                                     |                  |                 |
| GTZAN                   | 1000     | 389      | Numeric | 10           | 100.000 $\pm$ 0.000                                 | Genre            | 30s             |
| HOMBURG                 | 1886     | 389      | Numeric | 9            | 209.556 $\pm$ 134.866                               | Genre            | 10s             |
| FMA_SMALL               | 7996     | 389      | Numeric | 8            | 999.500 $\pm$ 0.500                                 | Genre            | 30s             |
| MER500                  | 493      | 389      | Numeric | 5            | 98.600 $\pm$ 2.332                                  | Emotion          | 30s $\pm$ 30s   |
| emoMusic                | 1000     | 389      | Numeric | 8            | 125.000 $\pm$ 0.000                                 | Emotion          | 45s             |
| Soundtracks             | 470      | 389      | Numeric | 9            | 52.222 $\pm$ 19.876                                 | Emotion          | 24s $\pm$ 14s   |
| Emotifymusic            | 400      | 389      | Numeric | 4            | 100.000 $\pm$ 0.000                                 | Genre            | 43 $\pm$ 18s    |
| Ballroom                | 698      | 389      | Numeric | 8            | 87.250 $\pm$ 17.555                                 | Genre            | 30s             |
| SeyerlehnerUnique       | 3115     | 389      | Numeric | 10           | 311.500 $\pm$ 248.349                               | Genre            | 30s             |
| MIREX-like_mood         | 903      | 385      | Numeric | 28           | 32.250 $\pm$ 4.778                                  | Mood             | 30s             |
| Medley-solos            | -        | -        | Numeric | 8            | 2537.875 $\pm$ 2016.080                             | Instrument       | 3s              |
| ISMIR04                 | 727      | 389      | Numeric | 6            | 121.167 $\pm$ 95.602                                | Genre            | 30s             |
| good-sounds             | 8399     | 371      | Numeric | 5            | 549.400 $\pm$ 198.697                               | Instrument       | 9s $\pm$ 9s     |
| MICM                    | 1137     | 389      | Numeric | 7            | 162.429 $\pm$ 119.717                               | Modal System     | 6m $\pm$ 6m     |
| Tropical Genres         | 1500     | 389      | Numeric | 5            | 300.000 $\pm$ 0.000                                 | Genre            | 30s             |
| PCMIR                   | 2410     | 387      | Numeric | 6            | 401.667 $\pm$ 46.241                                | Instrument       | 5s              |
| extendedBallroom        | 4180     | 389      | Numeric | 13           | 321.538 $\pm$ 195.702                               | Genre            | 2m $\pm$ 2m     |
| giantsteps_tempo        | 664      | 389      | Numeric | 23           | 28.870 $\pm$ 31.872                                 | Tempo            | 1m30s $\pm$ 30s |
| giantsteps_key          | 604      | 389      | Numeric | 24           | 25.167 $\pm$ 20.418                                 | Key              | 1m20s $\pm$ 40s |
| GMD                     | 1042     | -        | Numeric | 18           | 57.889 $\pm$ 70.069                                 | Genre            | 5m $\pm$ 5m     |
| FMA_MEDIUM              | 24984    | 389      | Numeric | 16           | 1561.500 $\pm$ 2069.356                             | Genre            | 30s             |
| IRMAS                   | 6700     | -        | Numeric | 5            | -                                                   | Instrument       | 2s              |
| MagnaTagATune_emotion   | 25860    | -        | Numeric | 17           | -                                                   | Emotion          | 30s             |
| MagnaTagATune_genre     | 25860    | -        | Numeric | 58           | -                                                   | Genre            | 30s             |
| MagnaTagATune_top50     | 25860    | -        | Numeric | 45           | -                                                   | Instrument       | 30s             |
| MagnaTagATune           | 25860    | -        | Numeric | 9            | -                                                   | Tempo            | 30s             |
| CAL500_emotion          | 498      | -        | Numeric | 27           | -                                                   | Emotion          | 11m $\pm$ 11m   |
| CAL500_instrument&vocal | 482      | -        | Numeric | 14           | -                                                   | Instrument&Vocal | 11m $\pm$ 11m   |
| CAL500_genre            | 449      | -        | Numeric | 12           | -                                                   | Genre            | 11m $\pm$ 11m   |

# Experimental settings(KNN/NB)-900 features

| Dataset                 | KNN(standardize=1)     | NB + laim              |
|-------------------------|------------------------|------------------------|
| GTZAN                   | 0.5980 ± 0.0302        | <b>0.7120 ± 0.0218</b> |
| HOMBURG                 | 0.2133 ± 0.0244        | <b>0.2761 ± 0.0155</b> |
| FMA_SMALL               | <b>0.2193 ± 0.0085</b> | 0.1882 ± 0.0110        |
| MER500                  | 0.4949 ± 0.0417        | <b>0.7214 ± 0.0363</b> |
| emoMusic                | 0.1610 ± 0.0288        | <b>0.2970 ± 0.0193</b> |
| Soundtracks             | 0.1528 ± 0.0365        | <b>0.4306 ± 0.0477</b> |
| Emotifymusic            | 0.6250 ± 0.0559        | <b>0.7275 ± 0.0381</b> |
| Ballroom                | 0.5065 ± 0.0324        | <b>0.7525 ± 0.0451</b> |
| SeyerlehnerUnique       | <b>0.6116 ± 0.0194</b> | 0.6031 ± 0.0152        |
| MIREX-like_mood         | 0.0583 ± 0.0142        | <b>0.1728 ± 0.0335</b> |
| Medley-solos            | <b>0.9164 ± 0.0057</b> | 0.7484 ± 0.0030        |
| ISMIR04                 | -                      | -                      |
| good-sounds             | <b>0.1068 ± 0.0053</b> | 0.0833 ± 0.0073        |
| MICM                    | 0.0714 ± 0.0149        | 0.0758 ± 0.0205        |
| Tropical Genres         | 0.9130 ± 0.0190        | <b>0.9160 ± 0.0102</b> |
| PCMR                    | 0.7618 ± 0.0183        | <b>0.7772 ± 0.0163</b> |
| extendedBallroom        | 0.3754 ± 0.1143        | <b>0.5311 ± 0.0199</b> |
| giantsteps_tempo        | <b>0.1788 ± 0.0301</b> | 0.0523 ± 0.0610        |
| giantsteps_key          | 0.0625 ± 0.0201        | <b>0.1267 ± 0.0500</b> |
| GMD                     | <b>0.2139 ± 0.0144</b> | 0.1851 ± 0.0261        |
| FMA_MEDIUM              | <b>0.5426 ± 0.0058</b> | 0.4653 ± 0.0075        |
| IRMAS                   | -                      | -                      |
| MagnaTagATune_emotion   | -                      | -                      |
| MagnaTagATune_genre     | -                      | -                      |
| MagnaTagATune_top50     | -                      | -                      |
| MagnaTagATune           | -                      | -                      |
| CAL500_emotion          | -                      | -                      |
| CAL500_instrument&vocal | -                      | -                      |
| CAL500_genre            | -                      | -                      |

- Overall Settings
  - 8:2 hold-out cross validation
  - “NaN” values were treated by KNN Imputer
  - Datasets consists of 900 under features
  - Experiment Iterations = 10
- KNN settings
  - KNN classifier (k=3)
- NB settings
  - Numeric type of datasets were dicretized by using laim discretization algorithm

# Experimental settings(KNN/NB)-400 features

| Dataset                 | KNN(standardize)       | NB + laim              |
|-------------------------|------------------------|------------------------|
| GTZAN                   | 0.5980 ± 0.0304        | <b>0.7220 ± 0.0270</b> |
| HOMBURG                 | 0.4194 ± 0.0231        | <b>0.4645 ± 0.0243</b> |
| FMA_SMALL               | 0.4621 ± 0.0150        | <b>0.4716 ± 0.0107</b> |
| MER500                  | 0.4582 ± 0.0286        | <b>0.7837 ± 0.0339</b> |
| emoMusic                | 0.2105 ± 0.0332        | <b>0.3325 ± 0.0399</b> |
| Soundtracks             | 0.2223 ± 0.0390        | <b>0.3436 ± 0.0420</b> |
| Emotifymusic            | 0.6388 ± 0.0356        | <b>0.7100 ± 0.0337</b> |
| Ballroom                | 0.5129 ± 0.0408        | <b>0.8158 ± 0.0212</b> |
| SeyerlehnerUnique       | 0.6019 ± 0.0168        | <b>0.6504 ± 0.0197</b> |
| MIREX-like_mood         | 0.0567 ± 0.0110        | <b>0.1828 ± 0.0289</b> |
| Medley-solos            | <b>0.8985 ± 0.0056</b> | 0.7665 ± 0.0063        |
| ISMIR04                 | <b>0.7814 ± 0.0309</b> | 0.7593 ± 0.0235        |
| good-sounds             | <b>0.1070 ± 0.0054</b> | 0.0818 ± 0.0055        |
| MICM                    | 0.0678 ± 0.0140        | <b>0.0996 ± 0.0321</b> |
| Tropical Genres         | 0.9290 ± 0.0187        | <b>0.9317 ± 0.0149</b> |
| PCMR                    | 0.7089 ± 0.0207        | <b>0.7938 ± 0.0101</b> |
| extendedBallroom        | 0.5812 ± 0.0169        | <b>0.7898 ± 0.0145</b> |
| giantsteps_tempo        | 0.1621 ± 0.0650        | <b>0.1909 ± 0.0509</b> |
| giantsteps_key          | 0.0700 ± 0.0292        | <b>0.0817 ± 0.0584</b> |
| GMD                     | <b>0.2803 ± 0.0350</b> | 0.2500 ± 0.0572        |
| FMA_MEDIUM              | <b>0.5376 ± 0.0130</b> | 0.4860 ± 0.0045        |
| IRMAS                   |                        | -                      |
| MagnaTagATune_emotion   |                        | -                      |
| MagnaTagATune_genre     |                        | -                      |
| MagnaTagATune_top50     |                        | -                      |
| MagnaTagATune           |                        | -                      |
| CAL500_emotion          |                        | -                      |
| CAL500_instrument&vocal |                        | -                      |
| CAL500_genre            |                        | -                      |

- Overall Settings
  - 8:2 hold-out cross validation
  - “NaN” values were treated by KNN Imputer
  - Datasets consists of 900 under features
  - Experiment Iterations = 10
- KNN settings
  - KNN classifier (k=3)
- NB settings
  - Numeric type of datasets were dicretized by using laim discretization algorithm

# Experimental settings(proposed)-900 features

| Dataset                 | proposed            |
|-------------------------|---------------------|
| GTZAN                   | 0.7570 $\pm$ 0.0303 |
| HOMBURG                 | 0.3117 $\pm$ 0.0288 |
| FMA_SMALL               | 0.2014 $\pm$ 0.0156 |
| MER500                  | 0.7694 $\pm$ 0.0376 |
| emoMusic                | 0.2880 $\pm$ 0.0280 |
| Soundtracks             | 0.4542 $\pm$ 0.0516 |
| Emotifymusic            | 0.7425 $\pm$ 0.0438 |
| Ballroom                | 0.8367 $\pm$ 0.0464 |
| SeyerlehnerUnique       | 0.6559 $\pm$ 0.0192 |
| MIREX-like_mood         | 0.1700 $\pm$ 0.0257 |
| Medley-solos            | 0.7870 $\pm$ 0.0065 |
| ISMIR04                 | -                   |
| good-sounds             | 0.1999 $\pm$ 0.0061 |
| MICM                    | 0.0780 $\pm$ 0.0320 |
| Tropical Genres         | 0.9297 $\pm$ 0.0109 |
| PCMIR                   | 0.8201 $\pm$ 0.0159 |
| extendedBallroom        | 0.6349 $\pm$ 0.0141 |
| giantsteps_tempo        | 0.0523 $\pm$ 0.0610 |
| giantsteps_key          | 0.1375 $\pm$ 0.0450 |
| GMD                     | 0.2625 $\pm$ 0.0263 |
| FMA_MEDIUM              | 0.5198 $\pm$ 0.0088 |
| IRMAS                   | -                   |
| MagnaTagATune_emotion   | -                   |
| MagnaTagATune_genre     | -                   |
| MagnaTagATune_top50     | -                   |
| MagnaTagATune           | -                   |
| CAL500_emotion          | -                   |
| CAL500_instrument&vocal | -                   |
| CAL500_genre            | -                   |

- Overall Settings
  - 8:2 hold-out cross validation
  - “NaN” values were treated by KNN Imputer
  - Datasets consists of 900 under features
  - Experiment Iterations = 10
- NB settings
  - Numeric type of datasets were dicretized by using laim discretization algorithm
- Extract entropy
  - Number of iterations : 10
  - Each score consists of f\_ent(single feature entropy), fl\_ent(feature-label entropy), ff\_ent(feature-feature entropy)
- Proposed
  - Initial population : 50
  - FFCS : 300
  - Initialization : 101 Initialization
  - Evaluation : Naïve Bayes

# Comparison - 900 features

| Dataset                 | KNN(standardize=1) | KNN(cosine)            | NB + laim              | Proposed(knn + relief) | Proposed(nb + gics)    |
|-------------------------|--------------------|------------------------|------------------------|------------------------|------------------------|
| GTZAN                   | 0.5980 ± 0.0302    | 0.6385 ± 0.0285        | 0.7120 ± 0.0218        | 0.6505 ± 0.0194        | <b>0.7570 ± 0.0303</b> |
| HOMBURG                 | 0.2133 ± 0.0244    | 0.2077 ± 0.0297        | 0.2761 ± 0.0155        | 0.2011 ± 0.0289        | <b>0.3117 ± 0.0288</b> |
| FMA_SMALL               | 0.2193 ± 0.0085    | <b>0.2338 ± 0.0081</b> | 0.1882 ± 0.0110        | -                      | 0.2014 ± 0.0156        |
| MER500                  | 0.4949 ± 0.0417    | 0.5449 ± 0.0428        | 0.7214 ± 0.0363        | 0.5367 ± 0.0391        | <b>0.7694 ± 0.0376</b> |
| emoMusic                | 0.1610 ± 0.0288    | 0.1635 ± 0.0201        | <b>0.2970 ± 0.0193</b> | 0.1600 ± 0.0170        | 0.2880 ± 0.0280        |
| Soundtracks             | 0.1528 ± 0.0365    | 0.1972 ± 0.0485        | 0.4306 ± 0.0477        | 0.1806 ± 0.0404        | <b>0.4542 ± 0.0516</b> |
| Emotifymusic            | 0.6250 ± 0.0559    | 0.6975 ± 0.0510        | 0.7275 ± 0.0381        | 0.1600 ± 0.0170        | <b>0.7425 ± 0.0438</b> |
| Ballroom                | 0.5065 ± 0.0324    | 0.5583 ± 0.0327        | 0.7525 ± 0.0451        | 0.5799 ± 0.0394        | <b>0.8367 ± 0.0464</b> |
| SeyerlehnerUnique       | 0.6116 ± 0.0194    | <b>0.6595 ± 0.0156</b> | 0.6031 ± 0.0152        | 0.6542 ± 0.0234        | 0.6559 ± 0.0192        |
| MIREX-like_mood         | 0.0583 ± 0.0142    | 0.0667 ± 0.0176        | <b>0.1728 ± 0.0335</b> | 0.0667 ± 0.0117        | 0.1700 ± 0.0257        |
| Medley-solos            | 0.9164 ± 0.0057    | <b>0.9350 ± 0.0030</b> | 0.7484 ± 0.0030        | -                      | 0.7870 ± 0.0065        |
| ISMIR04                 | -                  | -                      | -                      | -                      | -                      |
| good-sounds             | 0.1068 ± 0.0053    | 0.1067 ± 0.0054        | 0.0833 ± 0.0073        | -                      | <b>0.1999 ± 0.0061</b> |
| MICM                    | 0.0714 ± 0.0149    | 0.0621 ± 0.0159        | 0.0758 ± 0.0205        | 0.0648 ± 0.0268        | <b>0.0780 ± 0.0320</b> |
| Tropical Genres         | 0.9130 ± 0.0190    | 0.9220 ± 0.0164        | 0.9160 ± 0.0102        | <b>0.9593 ± 0.0165</b> | 0.9297 ± 0.0109        |
| PCMIR                   | 0.7618 ± 0.0183    | 0.8193 ± 0.0184        | 0.7772 ± 0.0163        | <b>0.8662 ± 0.0175</b> | 0.8201 ± 0.0159        |
| extendedBallroom        | 0.3754 ± 0.1143    | 0.3959 ± 0.1242        | 0.5311 ± 0.0199        | -                      | <b>0.6349 ± 0.0141</b> |
| giantsteps_tempo        | 0.1788 ± 0.0301    | 0.1902 ± 0.0375        | 0.0523 ± 0.0610        | <b>0.1955 ± 0.0278</b> | 0.0523 ± 0.0610        |
| giantsteps_key          | 0.0625 ± 0.0201    | 0.0625 ± 0.0249        | 0.1267 ± 0.0500        | 0.0567 ± 0.0235        | <b>0.1375 ± 0.0450</b> |
| GMD                     | 0.2139 ± 0.0144    | 0.2212 ± 0.0181        | 0.1851 ± 0.0261        | 0.2115 ± 0.0213        | <b>0.2625 ± 0.0263</b> |
| FMA_MEDIUM              | 0.5426 ± 0.0058    | <b>0.5733 ± 0.0059</b> | 0.4653 ± 0.0075        | -                      | 0.5198 ± 0.0088        |
| IRMAS                   | -                  | -                      | -                      | -                      | -                      |
| MagnaTagATune_emotion   | -                  | -                      | -                      | -                      | -                      |
| MagnaTagATune_genre     | -                  | -                      | -                      | -                      | -                      |
| MagnaTagATune_top50     | -                  | -                      | -                      | -                      | -                      |
| MagnaTagATune           | -                  | -                      | -                      | -                      | -                      |
| CAL500_emotion          | -                  | -                      | -                      | -                      | -                      |
| CAL500_instrument&vocal | -                  | -                      | -                      | -                      | -                      |
| CAL500_genre            | -                  | -                      | -                      | -                      | -                      |

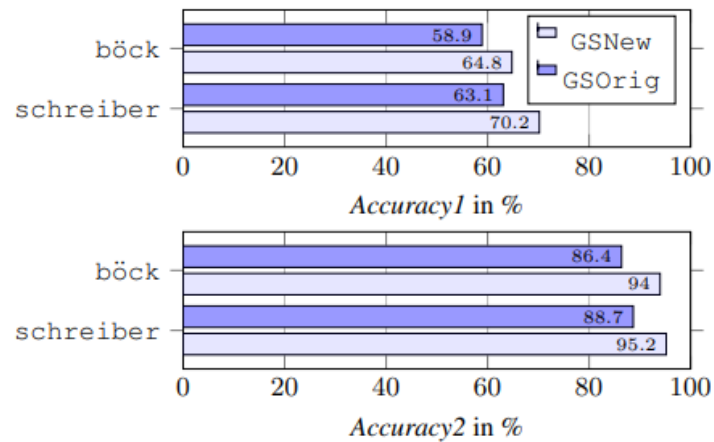
# Accuracies in other papers

## 1. MICM

TABLE III. DNN RESULTS WITHOUT GRU AND BOTTELNECH

| Name of Dastgah | Precision | Recall | F1    |
|-----------------|-----------|--------|-------|
| Shour           | 96.35     | 85.94  | 90.85 |
| Homayoun        | 75.34     | 81.28  | 78.20 |
| Mahour          | 90.42     | 87.12  | 88.74 |
| Segah           | 81.79     | 86.16  | 83.92 |
| Chahargah       | 62.23     | 86.57  | 72.41 |
| Rastpanjgah     | 88.06     | 90.74  | 89.38 |
| Nava            | 94.57     | 86.10  | 90.14 |

## 2. Giantsteps\_tempo



**Figure 9:** Accuracies for the algorithms böck and schreiber measured against both GSOrig and GSNew.

# References

---

[1] GTZAN

Liu, C., Feng, L., Liu, G. et al. *Bottom-up broadcast neural network for music genre classification*. *Multimed Tools Appl* 80, 7313–7331 (2021). <https://doi.org/10.1007/s11042-020-09643-6>

[2] HOMBURG

Medhat, F., Chesmore, D., & Robinson, J. (2020). *Masked Conditional Neural Networks for sound classification*. *Applied Soft Computing*, 106073. doi:10.1016/j.asoc.2020.106073

[3] FMA\_SMALL

Park, J., Lee, J., Park, J., Ha, J., & Nam, J. (2018). *Representation Learning of Music Using Artist Labels*. 19th International Society for Music Information Retrieval Conference (ISMIR), 2018

[4] MER500

Velankar, M., & Kulkarni, P. (2018). *MUSIC EMOTION RECOGNITION FOR INDIAN FILM MUSIC*. 19th International Society for Music Information Retrieval Conference (ISMIR), 2018

[5] emoMusic

Han, J., Zhang, Z., Ren, Z., & Schuller, B. (2020). *Exploring Perception Uncertainty for Emotion Recognition in Dyadic Conversation and Music Listening*. *Cognitive Computation*. doi:10.1007/s12559-019-09694-4

[6] Soundtracks

Er, M., & Aydılek, I. (2019). *Music Emotion Recognition by Using Chroma Spectrogram and Deep Visual Features*. IJCIS. Doi: <https://doi.org/10.2991/ijcis.d.191216.001>

[7] Emotifymusic

Jakubik, J. (2017). *Evaluation of Gated Recurrent Neural Networks in Music Classification Tasks*. *Advances in Intelligent Systems and Computing*, 27–37. doi:10.1007/978-3-319-67220-5\_3

[8] Ballroom

Liu, C., Feng, L., Liu, G. et al. *Bottom-up broadcast neural network for music genre classification*. *Multimed Tools Appl* 80, 7313–7331 (2021). <https://doi.org/10.1007/s11042-020-09643-6>

[9] SeyerlehnerUnique

Hamel, P., Davies, M., Yoshii, K. & Goto, M. (2013). *Transfer Learning In MIR: Sharing Learned Latent Representations For Music Audio Classification And Similarity*. 14th International Conference on Music Information Retrieval (ISMIR '13)

[10] MIREX-like\_mood

Bischoff, K., Firan, C., Paiu, R., & Nejd, W. (2009). *MUSIC MOOD AND THEME CLASSIFICATION - A HYBRID APPROACH*. 10th International Society for Music Information Retrieval Conference (ISMIR 2009)

[11] Medley-solos

Lagrange, M., & Gontier, F. (2020). *Bandwidth Extension of Musical Audio Signals With No Side Information Using Dilated Convolutional Neural Networks*. *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. doi:10.1109/icassp40776.2020.9054194

[12] ISMIR04

Medhat, F., Chesmore, D., & Robinson, J. (2020). *Masked Conditional Neural Networks for sound classification*. *Applied Soft Computing*, 106073. doi:10.1016/j.asoc.2020.106073

[13] good-sounds

Guizzo, E., Weyde, T., & Tarroni, G. (2021). *Anti-Transfer Learning for Task Invariance in Convolutional Neural Networks for Speech Processing*.

[14] MICM

Azar, S., Malekzadeh, S., Ahmadi, A., & Samami, M. (2018). *Instrument-Independent Dastgah Recognition of Iranian Classical Music Using AzarNet*. 27th Iranian Conference on Electrical Engineering (ICEE 2019)

# References

---

[15] Tropical Genres

[16] PCMIR

Mousavi, S. M. H., Prasath, V. B. S., & Mousavi, S. M. H. (2019). *Persian Classical Music Instrument Recognition (PCMIR) Using a Novel Persian Music Database*. 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE). doi:10.1109/iccke48569.2019.8965166

[17] extendedBallroom

Liu, C., Feng, L., Liu, G. et al. *Bottom-up broadcast neural network for music genre classification*. *Multimed Tools Appl* 80, 7313–7331 (2021). <https://doi.org/10.1007/s11042-020-09643-6>

[18] giantsteps\_tempo

Schreiber, H. & Muller, M. (2017). *A POST-PROCESSING PROCEDURE FOR IMPROVING MUSIC TEMPO ESTIMATES USING SUPERVISED LEARNING*. (ISMIR 2017)

[19] giantsteps\_key

Korzeniowski, F. & Widmer, G. (2018). *Genre-Agnostic Key Classification With Convolutional Neural Networks*. 19th International Society for Music Information Retrieval Conference

[20] GMD

Callender, L., Hawthorne, C., & Engel, J. (2020). *Improving Perceptual Quality of Drum Transcription with the Expanded Groove MIDI Dataset*.

[21] FMA\_MEDIUM

Boxler, D. (2015). *MACHINE LEARNING TECHNIQUES APPLIED TO MUSICAL GENRE RECOGNITION*.

[22] IRMAS

Pons, J., Slizovskaia, O., Gong, R., Gomez, E., & Serra, X. (2017). *Timbre analysis of music audio signals with convolutional neural networks*. 2017 25th European Signal Processing Conference (EUSIPCO). doi:10.23919/eusipco.2017.8081710

[23] MagnaTagATune\_emotion

Chen, Qinyue, and Adiba Orzikulova. *AUDIO-BASED MULTI-LABEL MOOD RECOGNITION USING DATA AUGMENTATION*.

[24] MagnaTagATune\_genre

Wolff, Daniel, and Tillman Weyde. *Adapting similarity on the magnatagatune database: effects of model and feature choices*. Proceedings of the 21st International Conference on World Wide Web. 2012.

[25] MagnaTagATune\_top50

[26] MagnaTagATune

[27] CAL500\_emotion

Liu, X., Chen, Q., Wu, X., Liu, Y., & Liu, Y. (2017). *CNN based music emotion classification*

[28] CAL500\_instrument&vocal

Wang, S.-Y., Wang, J.-C., Yang, Y.-H., & Wang, H.-M. (2014). *Towards time-varying music auto-tagging based on CAL500 expansion*. 2014 IEEE International Conference on Multimedia and Expo (ICME). doi:10.1109/icme.2014.6890290

[29] CAL500\_genre

Bodo, Z. & Szilagyi, E. (2018). *Connecting the Last.fm Dataset to LyricWiki and MusicBrainz*. Lyrics-based experiments in genre classification . Acta Universitatis Sapientiae, Informatica

# Appendix

# A review of population initialization techniques for EA

Main goal of those studies are

1. Increase the probability of finding global optima
2. Reduce the variation of final searching results
3. Decrease the computational costs
4. Improve the solution quality

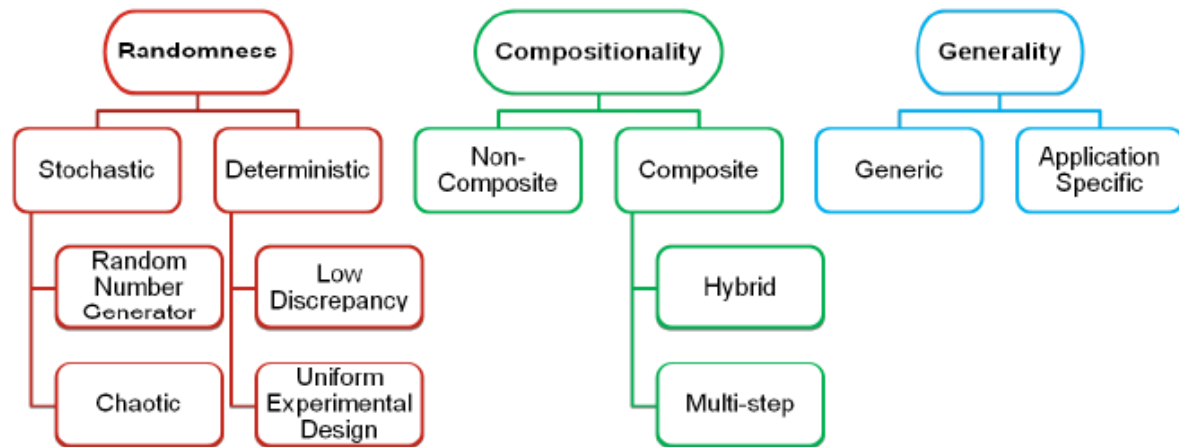


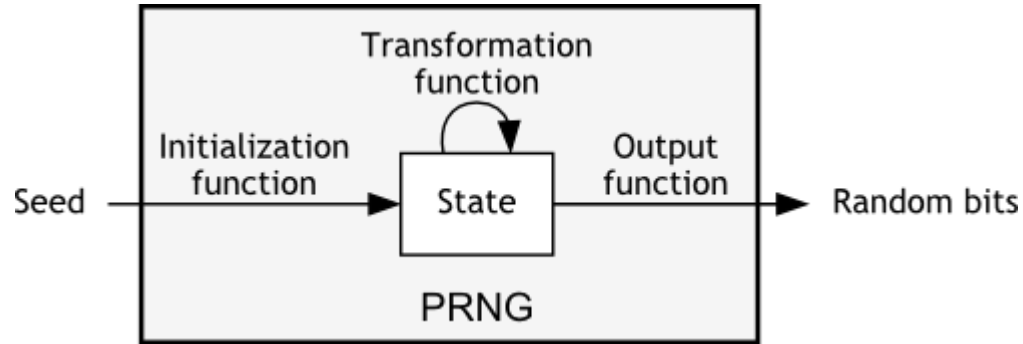
Fig. 1. Three categorizations of population initialization techniques, based on randomness, compositionality and generality.

# A review of population initialization techniques for EA

## Randomness

### Stochastic Technique

- Pseudo-Random Number Generator (PRNG)



- Chaotic Number Generators (CNG)
  - Improve performance in terms of population diversity, success rate, convergence speed

### Deterministic Technique(low-discrepancy techniques)

- Quasi-Random Sequence (QRS)
  - Discrepancy calculation > running several EA runs
- Uniform Experimental Design (UED)

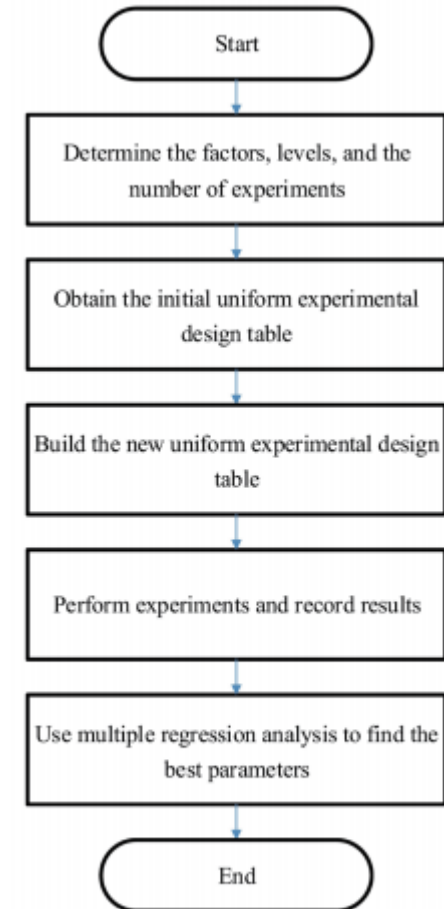


Fig. 2. Flow chart of the UED method.

# A review of population initialization techniques for EA

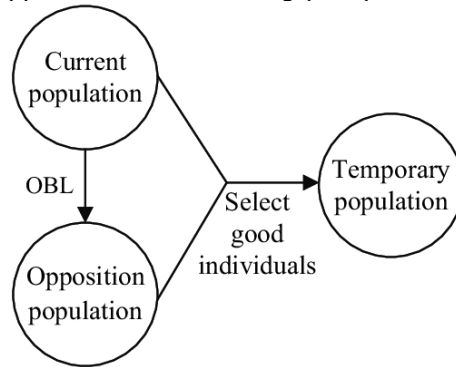
## Compositionality

### Non-compositional

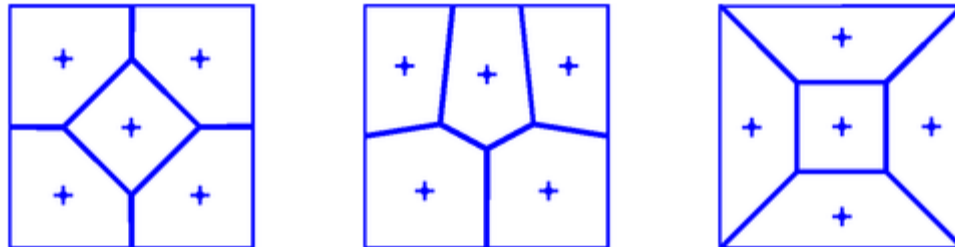
- Only one single step

### Compositional

- Comprise more than one stage
- Hybrid : Separately applied as a non-compositional technique
- Multi-step techniques : At least one of techniques cannot be employed as a standalone initializer
  - Opposition based learning (OBL)



- Centroidal Voronoi Tessellation (CVT)



# **A review of population initialization techniques for EA**

---

## **Generality**

- **Generic**
  - Can be directly applied to all types of optimization problems
  - No specific knowledge about the region of interest or building blocks of potential solution is available before running the EA
- **Application Specific**
  - Specially designed to be applied to particular real-world problems