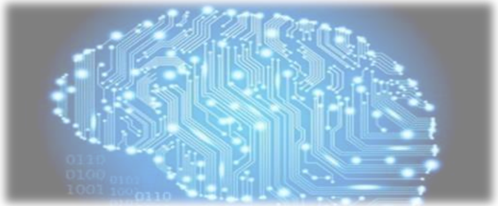
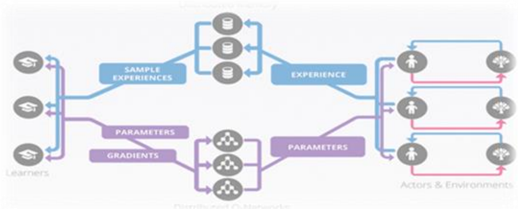




Statistics and results from KNN classifier and NB classifier

AutoML 연구실
석사 과정 조성현

21.05.11

Dataset statistics of single label - 900 features

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

Dataset statistics of single label - 400 features

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

Experimental settings(KNN/NB)-900 features

Dataset	Avg. of the results(KNN)	Avg. of the results(NB+laim)
GTZAN	**0.8564 ± 0.0095**	0.8552 ± 0.0128
HOMBURG	**0.7610 ± 0.0129**	0.6846 ± 0.0106
FMA_SMALL	0.8341 ± 0.0039	**0.8366 ± 0.0033**
MER500	**0.7131 ± 0.0158**	0.6873 ± 0.0265
emoMusic	**0.8619 ± 0.0049**	0.8487 ± 0.0069
Soundtracks	0.9012 ± 0.0114	**0.9022 ± 0.0122**
Emotifymusic	**0.6375 ± 0.0255**	0.6203 ± 0.0094
Ballroom	**0.7698 ± 0.0225**	0.7671 ± 0.0155
SeyerlehnerUnique	**0.9268 ± 0.0011**	0.9219 ± 0.0022
MIREX-like_mood	**0.9638 ± 0.0025**	0.9626 ± 0.0033
Medley-solos	0.6626 ± 0.0054	**0.6721 ± 0.0034**
ISMIR04	-	-
good-sounds	**0.9167 ± 0.0000**	**0.9167 ± 0.0000**
MICM	**0.8295 ± 0.0100**	0.8092 ± 0.0067
Tropical Genres	**0.7080 ± 0.0115**	0.6704 ± 0.0153
PCMIR	**0.7630 ± 0.0070**	0.7487 ± 0.0144
extendedBallroom	**0.9231 ± 0.0000**	0.9228 ± 0.0005
giantsteps_tempo	0.9033 ± 0.0117	**0.9054 ± 0.0178**
giantsteps_key	**0.9561 ± 0.0035**	0.9510 ± 0.0076
GMD	**0.9337 ± 0.0041**	0.9157 ± 0.0076
FMA_MEDIUM	**0.9363 ± 0.0004**	0.9336 ± 0.0008
IRMAS	-	-
MagnaTagATune_emotion	-	-
MagnaTagATune_genre	-	-
MagnaTagATune_instrument	-	-
MagnaTagATune_tempo	-	-
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
- 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	Avg. of the results(KNN)	Avg. of the results(NB+laim)
GTZAN	**0.8604 ± 0.0162**	0.8584 ± 0.0091
HOMBURG	**0.7170 ± 0.0172**	0.6026 ± 0.0090
FMA_SMALL	**0.8109 ± 0.0048**	0.7929 ± 0.0065
MER500	**0.6996 ± 0.0176**	0.6763 ± 0.0235
emoMusic	0.8403 ± 0.0159	**0.8492 ± 0.0062**
Soundtracks	**0.8383 ± 0.0192**	0.7997 ± 0.0283
Emotifymusic	**0.6438 ± 0.0169**	0.6188 ± 0.0161
Ballroom	**0.7678 ± 0.0222**	0.7637 ± 0.0100
SeyerlehnerUnique	**0.8960 ± 0.0020**	0.8901 ± 0.0036
MIREX-like_mood	0.9585 ± 0.0041	**0.9638 ± 0.0021**
Medley-solos	-	-
ISMIR04	**0.7659 ± 0.0111**	0.7613 ± 0.0138
good-sounds	-	-
MICM	0.8303 ± 0.0059	**0.8204 ± 0.0127**
Tropical Genres	**0.7048 ± 0.0072**	0.6800 ± 0.0170
PCMIR	**0.7623 ± 0.0135**	0.7462 ± 0.0045
extendedBallroom	**0.9231 ± 0.0000**	**0.9231 ± 0.0000**
giantsteps_tempo	**0.9030 ± 0.0143**	0.8831 ± 0.0215
giantsteps_key	0.9536 ± 0.0059	**0.9558 ± 0.0049**
GMD	-	-
FMA_MEDIUM	**0.9362 ± 0.0005**	0.9342 ± 0.0006
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
- KNN settings
  - KNN classifier (k=3)
- NB settings
  - Numeric type of datasets were dicretized by using laim discretization algorithm

Experimental settings(Proposed)

- Extract score
  - Number of iterations : 10
  - Each score consists of f_ent, fl_ent, ff_ent

- Proposed
  - Initial population : 50
  - FFCS : 300
  - Initialization : 101 Initialization
  - Evaluation : Naïve Bayes

- Classification accuracy

Iteration	Acc
1	0.2250
2	0.1875
3	0.2250
4	0.1875
5	0.2125
6	0.1750
7	0.2000
8	0.2000
9	0.2375
10	0.1875
Avg.	0.2038

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., & Aydilek, 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., & Nejdl, 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

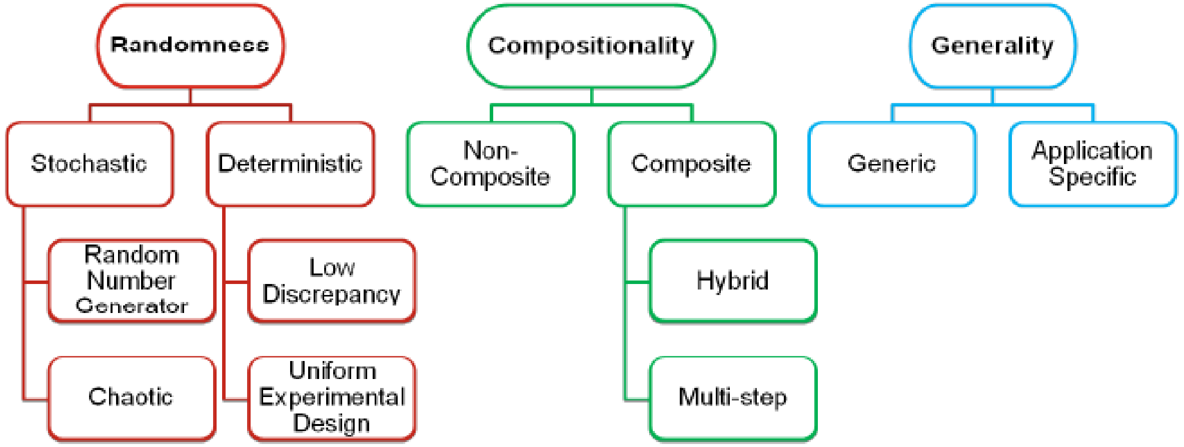


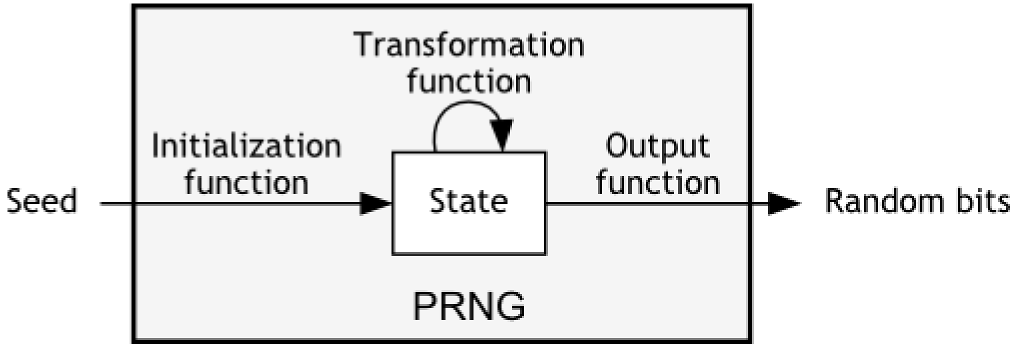
Fig. 1. Three categorizations of population initialization techniques, based on randomness, compositionality and generality.

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)

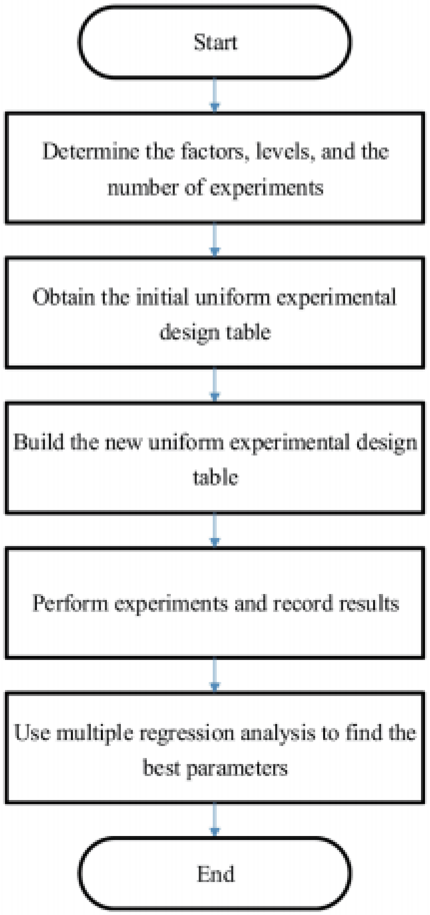


Fig. 2. Flow chart of the UED method.

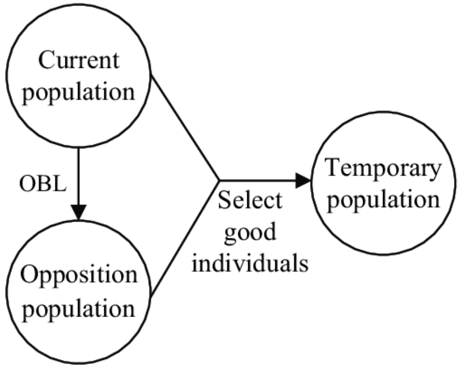
A review of population initialization techniques for EA

Compositionality

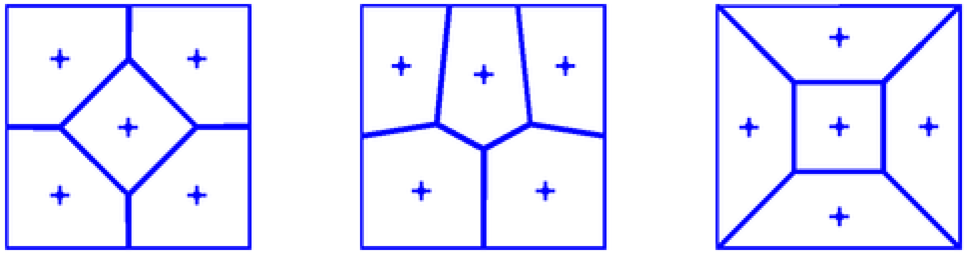
Non-compositional
- Only one single step

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