Generative Models Based on the Bounded Asymmetric Gaussian Distribution

Zixiang Xian [1]

Supervisors: Dr. Nizar Bouguila [1]
Dr. Manar Amayri [1,2]

[1] *Concordia Institute for Information Systems Engineering Concordia University, Montreal, Canada*

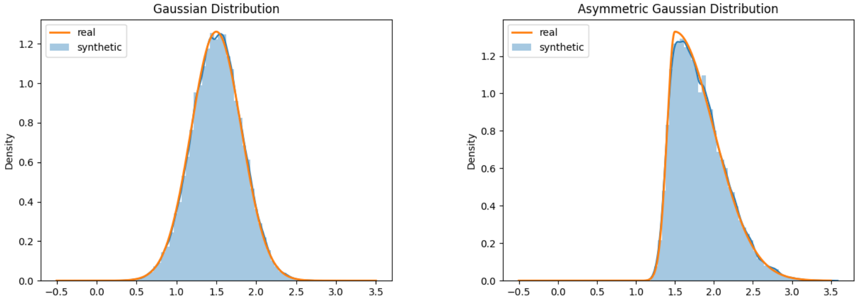
[2] *Grenoble Institute of Technology, Grenoble, France*

August 22, 2021

Outline

1. **Proposed Model**
   - Bounded Asymmetric Gaussian Distribution
   - Generative models based on BAGD
2. **Model Selection**
   - MML
   - Experimental Results
3. **Feature Selection**
   - Parameter Estimation
   - Model selection via MML
   - Human Activity Categorization
   - Gender Recognition
4. **Integration into HMM**
   - Hidden Markov Model
   - Occupancy Estimation
   - Human Activity Recognition (HAR)
5. **Conclusion**

Asymmetric Gaussian Distribution

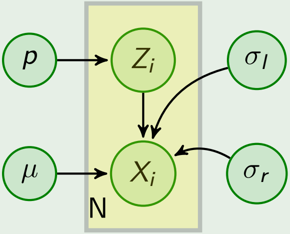


- Gaussian distribution (GD) assumes that the data is symmetric and has an infinite range, which prevents it from having a good modeling capability in the presence of outliers.
- Asymmetric Gaussian distribution (AGD) has been proposed to model asymmetric real-world data by having two variance parameters controlling the left and right parts of the distribution.

Proposed Model

Mixture of Asymmetric Gaussian Distributions (AGMM)

Graphical representation



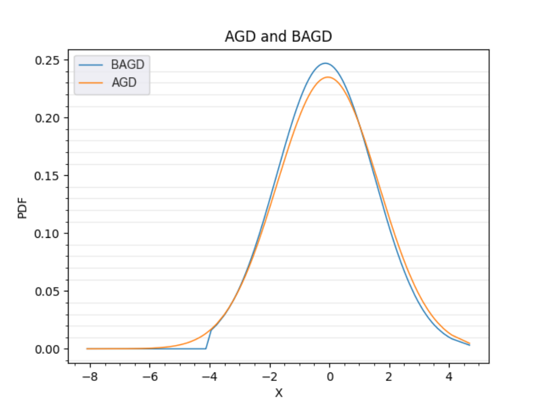
- p and Z_i are the mixing coefficient and posterior probability.
- $\mu, \sigma_l \& \sigma_r$ are the mean, left standard deviation and right standard deviation.

Mathmatical Definition

$$p(\mathcal{X}, \mathcal{Z}|\Theta) = \prod_{i=1}^{N}\prod_{j=1}^{K}\left(p(\vec{X}_i|\xi_j)p_j\right)^{Z_{ij}} \quad (1)$$

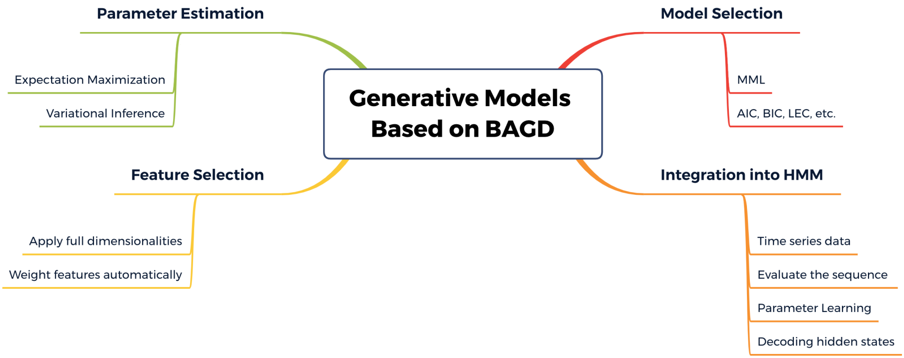
- $p\left(\vec{X}_i|\xi_j\right)$ is the PDF of AGD
- Z_{ij} is the hidden variable which satisfy $Z_{ij} \in \{0, 1\}$
- p_j are the mixing weights that satisfy $p_j \geq 0$, $\sum_{j=1}^{K} p_j = 1$.
- **Unbounded support.**

Mixture of Bounded Asymmetric Gaussian Distributions



$$p(\vec{X}|\xi_j) = \frac{f(\vec{X}|\xi_j)\mathrm{H}(\vec{X}|\Omega_j)}{\int_{\partial_j} f(\vec{u}|\xi_j)du}, \text{where } \mathrm{H}(\vec{X}|\Omega_j) = \begin{cases} 1 & \text{if } \vec{X} \in \partial_j \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$f(\vec{X}|\xi_j)$ is the PDF of AGD. The term $\int_{\partial_j} f(\vec{u}|\xi_j)du$ in Eq. (2) is the normalized constant that shows the share of $f(\vec{X}|\xi_j)$ which belongs to the support region ∂.



- ☞ Using EM algorithm to estimate the mixture's parameters.
- ☞ Adopting minimum message length (MML) as model selection criterion.

APPLICATIONS

- Occupancy Estimation
- Human Activity Recognition (HAR)
- Human Gender Recognition
- Documents Clustering
- Images/Videos Classification
- Named Entity Recognition (NER)
- Speech Recognition

Model selection criteria

EM algorithm requires an appropriate number of clusters found by model selection criteria.

- stochastic methods (e.g. Markov Chain Monte Carlo)
- deterministic approaches including Akaike's information criterion (AIC) [1], the Schwarzs Bayesian information criterion (BIC) [2], Consistent AIC (CAIC) [3], minimum description length (MDL) [4], the mixture minimum description length (MMDL) [5], the Laplace empirical criterion (LEC) [6] and minimum message length (MML) [7, 8].

The MML has been shown to have better performance among most model selection criteria in most cases for containing both prior distribution and the Fisher information matrix.

Minimum Message Length (MML)

$$\text{Mess Len}(K) \simeq -\log\left(p\left(\Theta_K\right)\right) - \mathcal{L}\left(\Theta_K, Z, \mathcal{X}\right) + \frac{1}{2}\log\left|F\left(\Theta_K\right)\right| + \frac{N_p}{2}\left(1 + \log\left(\frac{1}{12}\right)\right) \tag{3}$$

Θ_K	set of parameters when mixture contains K components
$p(\Theta_K)$	prior probability
$\mathcal{L}\left(\Theta_K, Z, \mathcal{X}\right)$	log-likelihood of mixture model
$\lvert F(\Theta_K)\rvert$	determinant of Fisher information matrix
N_p	number of parameters (equal to $K(3D+1)$)

Minimum Message Length (MML)

$$\text{Mess Len}(K) \simeq -\log\left(p\left(\Theta_K\right)\right) - \mathcal{L}\left(\Theta_K, Z, \mathcal{X}\right) + \frac{1}{2}\log\left|F\left(\Theta_K\right)\right| + \frac{N_p}{2}\left(1 + \log\left(\frac{1}{12}\right)\right) \quad (3)$$

$p(\Theta_K)$ prior probability $\Rightarrow$ Factorize prior distribution by assuming all parameters are mutually independent.

$$p(\Theta) = p(\pi)p(\mu)p(\sigma_l)p(\sigma_r) \quad (4)$$

$|F(\Theta_K)|$ determinant of Fisher information matrix $\Rightarrow$ Approximate the Hessian matrix by complete-data Fisher information matrix.

$$|F(\Theta)| = |F(\pi)||F(\mu)||F(\sigma_l)||F(\sigma_r)| \quad (5)$$

Model Learning Algorithm

Algorithm 1 Model Learning for BAGMM

Input:Dataset $\mathcal{X} = \{\vec{X}_1, \dots, \vec{X}_N\}$, t_{min}, K_{min}, K_{max}.
Output: Θ, $\mathcal{Z}$, K^*.
for $K_{min} \leq K \leq K_{max}$ do
  {**Initialization**}:
  K-Means (Compute $\vec{\mu}_1, \dots, \vec{\mu}_K$ & cluster assignment)
  for all $1 \leq j \leq K$ do
    Computation of p_j and $\{(\vec{\sigma}_{l_j}$ & $\vec{\sigma}_{r_j})=\vec{\sigma}_j\}$
  {**Expectation Maximization**}:
  while relative change in log-likelihood $\geq t_{min}$ **or** iterations $\leq epoch_{max}$ **or** relative changes of parameters $\geq t_{min}$ do
    {[**E Step**]}:
    for all $1 \leq j \leq K$ do
      Compute $p(j|\vec{X}_i)$ for $i = 1, \dots, N$.
    {[**M step**]}:
    update bounded support range
    for all $1 \leq j \leq K$ do
      Estimate p_j, $\vec{\mu}_j$, $\vec{\sigma}_{l_j}$ & $\vec{\sigma}_{r_j}$
  end while
  Compute $K^* = \arg\min MML(K)$
end for

Model Selection Results

Real Dataset											
dataset	N	D	K	AIC	BIC	CAIC	MDL	MMDL	MML like	LEC	MML
Indian Liver Patient(AGMM)	583	10	**2**	4	2	2	2	4	4	2	2
Indian Liver Patient(BAGMM)				2	2	2	2	2	2	**2**	**2**
Iris(AGMM)	150	4	**3**	6	3	3	3	3	6	6	6
Iris(BAGMM)				6	6	6	6	6	6	**3**	**3**
Vertebral(AGMM)	310	6	**3**	3	3	3	3	3	3	3	3
Vertebral(BAGMM)				5	3	3	3	5	5	**3**	**3**
Wine Quality(red)(AGMM)	1599	11	**6**	5	5	5	5	5	5	6	6
Wine Quality(red)(BAGMM)				8	8	8	8	8	8	**6**	**6**
Spect Heart(AGMM)	80	44	**2**	6	4	2	4	4	6	2	2
Spect Heart(BAGMM)				5	2	2	2	5	5	**2**	**2**
Cryotherapy(AGMM)	90	6	**2**	2	2	2	2	2	2	2	2
Cryotherapy(BAGMM)				6	2	2	2	6	6	**2**	**2**
Immunotherapy(AGMM)	90	7	**2**	3	2	2	2	3	3	2	2
Immunotherapy(BAGMM)				2	2	2	2	2	2	**2**	**2**
Statlog(Heart)(AGMM)	270	13	**2**	6	6	2	6	6	6	6	6
Statlog(Heart)(BAGMM)				2	2	2	2	2	2	**2**	**2**
Parkinsons(AGMM)	197	22	**2**	6	6	6	6	6	6	6	6
Parkinsons(BAGMM)				2	2	2	2	2	2	**2**	**2**
Haberman Survival(AGMM)	306	3	**2**	2	2	2	2	2	2	2	2
Haberman Survival(BAGMM)				2	2	2	2	2	2	**2**	**2**

Feature Selection

$$\log p(\mathcal{X}, Z \mid \Theta) = \sum_{i=1}^{N} \sum_{j=1}^{K} Z_{ij} \log \left[p\left(\vec{X}_i \mid \xi_j\right) p_j \right] \tag{6}$$

- All D features have **same weights** but input data consists of potentially irrelevant features.
- Irrelevant features can compromise the effectiveness of clustering and increase the computational complexity.
- Resulting in unreliable homogeneity measures.

$$p\left(\vec{X}_i \mid \Theta_K\right) = \sum_{j=1}^{K} p_j \prod_{d=1}^{D} \left[\omega_d p\left(X_d \mid \xi_{jd}\right) + (1-\omega_d)\, p\left(X_d \mid \lambda_d\right)\right] \tag{7}$$

Where $p\left(X_d \mid \lambda_d\right)$ is the background Gaussian distribution with the mean and standard deviation $\vec{\lambda} = \{\vec{\eta}, \vec{\delta}\}$. Assuming not all features have same relevancy by assigning weights to these features, denoted as $\vec{\omega} = (\omega_1, \cdots, \omega_D)$, where $0 \leq \omega_d \leq 1$, $d = 1, \ldots, D$ [7].

The log-likelihood function can be written as:

$$\begin{aligned}\mathcal{L}(\mathcal{X}, \mathcal{Z} \mid \Theta) &= \sum_{i=1}^{N} \sum_{j=1}^{K} Z_{ij} \log\left(p\left(\vec{X}_i \mid \Theta_K\right)\right) \\ &= \sum_{i=1}^{N} \sum_{j=1}^{K} Z_{ij} \left\{\log p_j + \log\left[\omega_d p\left(\vec{X}_i \mid \xi_j\right) + (1-\omega_d)\, p\left(\vec{X}_i \mid \lambda\right)\right]\right\}\end{aligned} \tag{8}$$

The parameters can be estimated by taking the gradient of the log-likelihood in the previous equation with respect to each parameters. The left and right standard deviation, $\sigma_{l_{jd}}$ & $\sigma_{r_{jd}}$, can be estimated using Newton-Raphson method.

Minimum Message Length

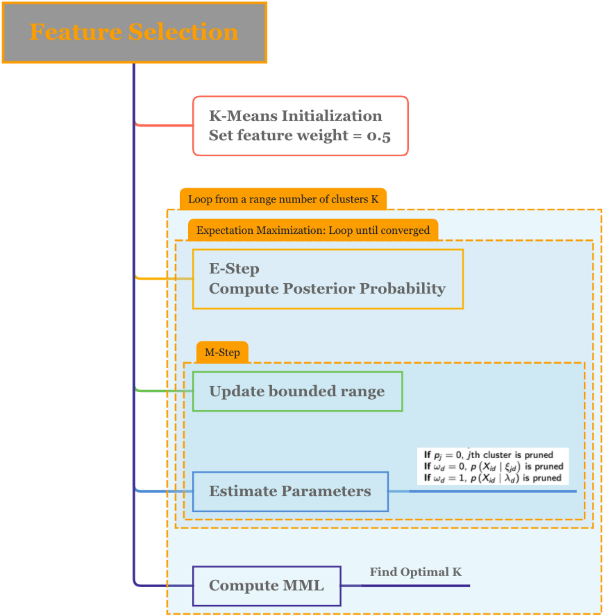
$$\text{MessLens} \approx -\log p\left(\Theta_M\right) + \frac{c}{2}\left(1 + \log\frac{1}{12}\right) + \frac{1}{2}\log\left|I\left(\Theta_M\right)\right| - \log p\left(\mathcal{X} \mid \Theta_M\right) \quad (9)$$

We assume that each group of parameters is independent, which allows the factorization of $p\left(\Theta_M\right)$ and $I\left(\Theta_M\right)$.

MML can be rewritten as

$$\begin{aligned}\text{MessLens} \approx &\ \frac{c}{2}\left(1 + \log\frac{1}{12}\right) + \frac{c}{2}(\log N) + \frac{3M}{2}\sum_{d=1}^{D}\log\omega_d \\ &+ \frac{3D}{2}\sum_{j=1}^{M}\log p_j + \sum_{d=1}^{D}\log\left(1-\omega_d\right) - \log p\left(\mathcal{X} \mid \theta_M\right)\end{aligned} \quad (10)$$

Model selection via MML



Clustering Metrics

- Accuracy is computed as: $\left(\frac{TP+TN}{TP+TN+FP+FN}\right)$
- Precision is computed as: $\left(\frac{TP}{TP+FP}\right)$
- Recall is computed as: $\left(\frac{TP}{TP+FN}\right)$
- F1 Score is computed as: $2 * (precision * recall)/(precision + recall)$
- Silhouette score indicates the overlapping clusters with the range from -1 to 1, and 1 is the best value, -1 for the worst value, value near 0 indicates overlapping clusters.
- Classification entropy (CE) index indicates good clustering when it is low and poor clustering when it is high.
- G-mean 1, the geometric mean of precision and recall, which is computed as: $\sqrt{precision \times recall}$
- G-mean 2, the geometric mean of specificity and recall, which is computed as: $\sqrt{specificity \times recall}$
- Mathew' s correlation coefficient (MCC) is computed as: $\frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$

the term TP stands for true positives, TN for true negatives, FP for false positives, and FN stands for false negatives

Human Activity Categorization[1]

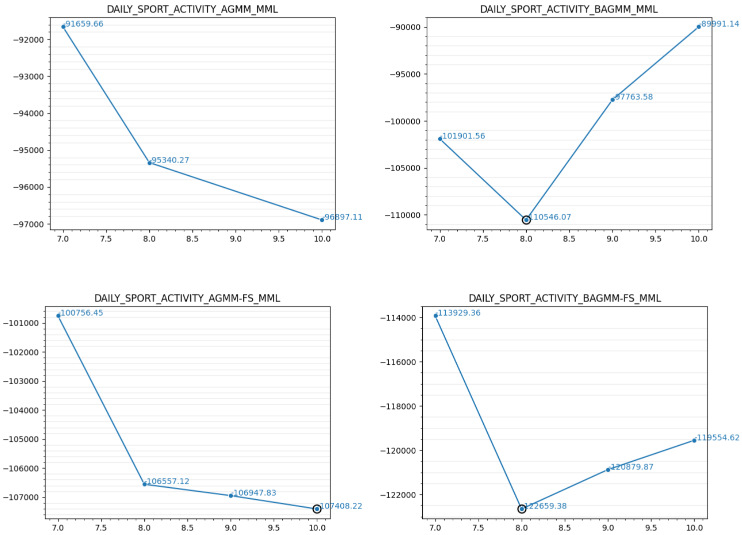
We consider a human activity categorization dataset called UCI Daily and Sports Activity dataset (DSAD)[1], which contains 19 different kinds of signal data performed by eight subjects (4 female, 4 male)
Eight daily activities from the first subject, including sitting, standing, walking, jumping, playing basketball, rowing, exercising, and running, are chosen to be classified.

Table: 8 common daily activities, from the first subject, clustering using different mixture models.

Models	Time	Epoch	Accuracy	Precision	Recall	F1-score	Silhouette	CE
BAGMM-FS	3.11	3	96.47%	97.25%	96.47%	96.40%	0.451	0.004
BAGMM	3.04	1	95.56%	96.33%	95.56%	95.29%	0.454	1.49
AGMM-FS	2.67	38	96.37%	97.18%	96.37%	96.29%	0.451	0.002
AGMM	0.468	12	95.86%	96.89%	95.86%	95.75%	0.452	0.002
BGGMM	494.95	7	95.56%	96.33%	95.56%	95.30%	0.454	4.17e-7
GGMM	0.640	7	62.5%	43.81%	62.50%	50.03%	0.198	0.430
GMM	0.014	1	44.76%	36.74%	44.75%	34.83%	0.211	0.641

[1] **DSAD** dataset available at: `http://archive.ics.uci.edu/ml/datasets/Daily+and+Sports+Activities`

Human Activity Categorization



Only BAGMM and BAGMM-FS were able to find the correct number of components which is 8, while AGMM and AGMM-FS favored 10 clusters.

Human Activity Categorization[2]

We also cluster different sitting activities from the 8 subjects in this dataset. **Feature selection improves the clustering results.** Mixture models without feature selection have almost the same accuracy as the baseline GMM. Our proposed model distinguishes itself as compared to the other mixture models with respect to all considered clustering metrics.

Table: Clustering of the sitting activities of the 8 subjects using different mixture models.

Models	Time	Epoch	Accuracy	Precision	Recall	F1-score	Silhouette	CE
BAGMM-FS	5.87	6	90.52%	93.13%	90.52%	89.80%	0.514	0.009
BAGMM	2.23	2	72.78%	70.88%	72.78%	71.47%	0.569	0.011
AGMM-FS	0.641	9	84.97%	78.35%	84.97%	80.43%	0.593	0.156
AGMM	0.105	1	72.47%	63.37%	72.47%	65.58%	0.624	0.384
BGGMM	107.74	16	72.47%	71.41%	72.48%	71.50%	0.495	4.4e-77
GGMM	0.237	3	72.47%	63.37%	72.47%	65.58%	0.624	0.384
GMM	0.015	1	71.67%	59.87%	71.67%	63.67%	0.556	0.383

We verify BAGMM-FS on three well-known datasets, PARSE-27k dataset (P27K), PETA dataset and Human attribute dataset (H3D).
Using bag of visual words (BOVW) to describe the images:

1. extract local features for each image using scale invariant feature transform (SIFT)
2. apply K-Means to cluster the 128-dimensional descriptors for building the visual words vocabulary

Figure: Samples images from datasets.

Table: Gender recognition results.

Models	dataset	Time	Epoch	Acc	Precision	Recall	F1-score	Silhouette	CE
BAGMM-FS	PETA	2.18	7	**81.21**	81.52%	81.21%	81.29%	0.018	0.170
BAGMM	PETA	1.322	4	51.44	48.73%	51.44%	48.97%	-0.002	0.011
AGMM-FS	PETA	0.813	45	57.80	33.41%	57.80%	42.34%	N/A	0.693
AGMM	PETA	0.237	21	57.80	33.41%	57.80%	42.34%	N/A	0.693
BGGMM	PETA	15.93	3	39.59	41.91%	39.59%	36.31%	0.075	0.011
GGMM	PETA	2.046	300	57.80	33.41%	57.80%	42.34%	N/A	0.693
GMM	PETA	0.024	1	57.80	33.41%	57.80%	42.34%	N/A	0.693
BAGMM-FS	P27K	3.13	5	**77.33**	82.49%	77.33%	67.83%	-0.122	0.005
BAGMM	P27K	2.02	4	50.18	82.47%	50.18%	51.47%	0.055	0.039
AGMM-FS	P27K	4.10	13	76.93	59.19%	76.93%	66.90%	N/A	0.693
AGMM	P27K	37.61	209	76.93	59.19%	76.93%	66.90%	N/A	0.693
BGGMM	P27K	30.33	6	70.61	76.80%	70.61%	72.52%	0.012	0.117
GGMM	P27K	1.101	3	76.93	59.19%	76.93%	66.90%	N/A	0.693
GMM	P27K	0.112	1	76.93	59.19%	76.93%	66.90%	N/A	0.693
BAGMM-FS	H3D	0.506	1	**70.61**	73.89%	70.61%	69.56%	0.125	0.116
BAGMM	H3D	0.504	1	62.77	78.66%	62.77%	56.79%	0.110	0.007
AGMM-FS	H3D	0.571	16	50.00	25.00%	50%	33.33%	N/A	0.693
AGMM	H3D	4.003	300	50.00	25.00%	50%	33.33%	N/A	0.693
BGGMM	H3D	48.073	8	61.98	62.43%	61.98%	61.62%	0.097	0.009
GGMM	H3D	0.121	3	50.00	25.00%	50%	33.33%	N/A	0.693
GMM	H3D	0.038	1	50.00	25.00%	50%	33.33%	N/A	0.693

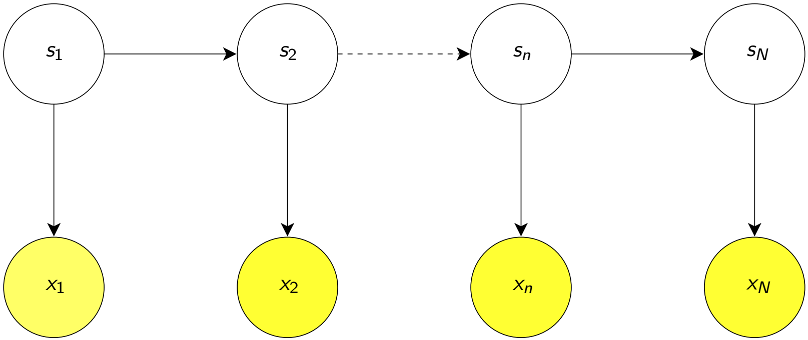
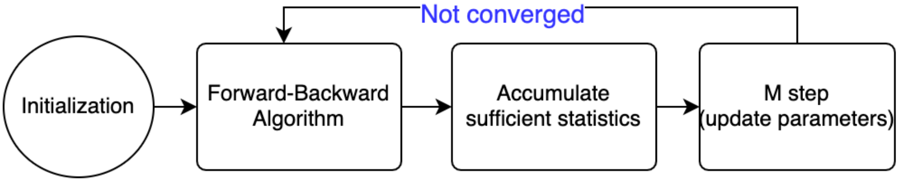


Figure: Graphical representation for HMM

$$p(\mathbf{X}, \mathbf{S} \mid \boldsymbol{\Theta}) = p\left(\mathbf{s}_1 \mid \boldsymbol{\pi}\right) \left[\prod_{n=2}^{N} p\left(\mathbf{s}_n \mid \mathbf{s}_{n-1}, \mathbf{A}\right)\right] \prod_{m=1}^{N} p\left(\vec{x}_m \mid \Lambda\right) \tag{11}$$

where $p\left(\vec{x}_m \mid \Lambda\right)$ is known as emission probability, which is bounded asymmetric Gaussian mixture in our paper.

Complete algorithm



Motivation for BAGMM Integration into HMM

- The bounded range support from BAGMMM and its asymmetric nature for modeling non-symmetric real-world data.
- Converge faster with less iterations and less time consuming.
- Efficiency in dealing with characterizing real-world signals based on time series.

Experiment for occupancy estimation

The dataset consists of environmental sensors data collected in an office in Grenoble Institute of Technology over a time interval $t = 30$ minutes with 5 hidden states $S = \{s_0, s_1, s_2, s_3, s_4\}$.

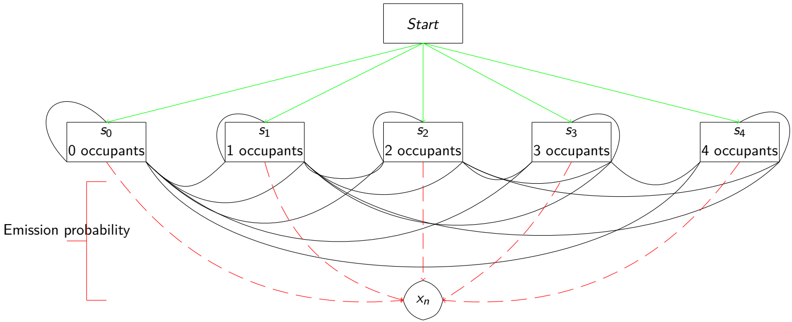
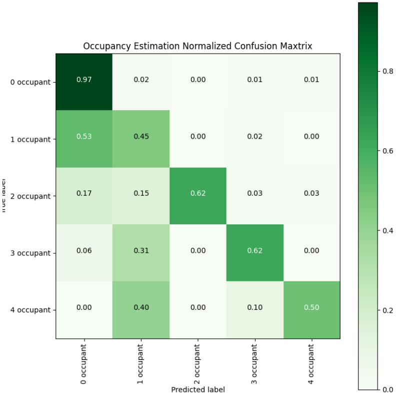


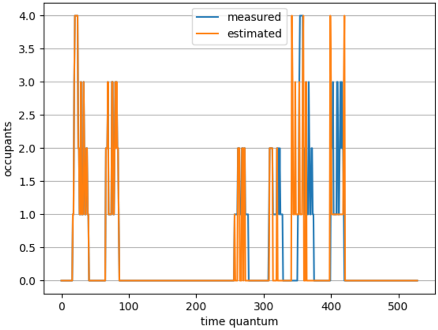
Figure: HMM for occupancy estimation accodring to the case of study

Table: Occupancy estimation comparison using different HMM models

	HMM Models			
Metrics	BAGMM-HMM	AGMM-HMM	BGMM-HMM	GMM-HMM
Epoch	4	4	2	10
Accuracy	86.39%	78.45%	75.42%	70.69%
Precision	85.71%	82.91%	56.89%	83.97%
Recall	86.38%	78.45%	75.42%	70.69%
Specificity	75.04%	82.47%	24.57%	88.57%
F1-score	85.52%	79.28%	64.86%	75.42%
G-mean 1	86.05%	80.66%	65.51%	77.05%
G-mean 2	80.52%	80.43%	43.05%	79.13%
MCC	68.35%	57.37%	52.28%	54.39%

In Table 5, the BAGMM-HMM achieves the best performance with an average accuracy of 86.39% and the highest F1-score with 85.52% compared to 78.45% and 79.28% for AGMM-HMM, 75.42%, and 64.86% for BGMM-HMM against 70.69%, and 75.42% for GMM-HMM, respectively.



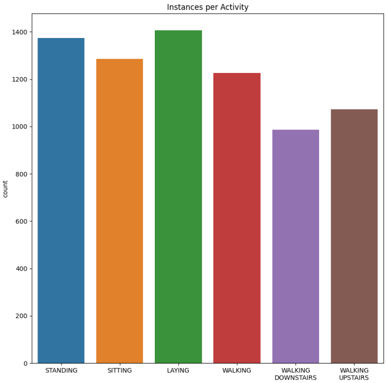


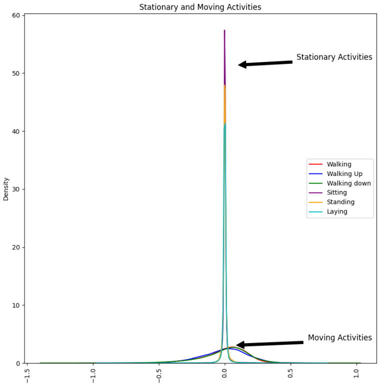
The dataset is an imbalanced dataset from the confusion matrix. But overall, our model can outperform the other HMM models with the same training data. BAGMM-HMM achieves 86.39% accuracy, compared with the ground truth as shown with the blue line.

Preprocessing and Data Visualization

challenging HAR dataset from UCI machine learning repository

There are two categorical activities: static (sitting, standing, laying) and dynamic (walking, walking upstairs, walking downstairs) activities according to body acceleration features in the y-axis.





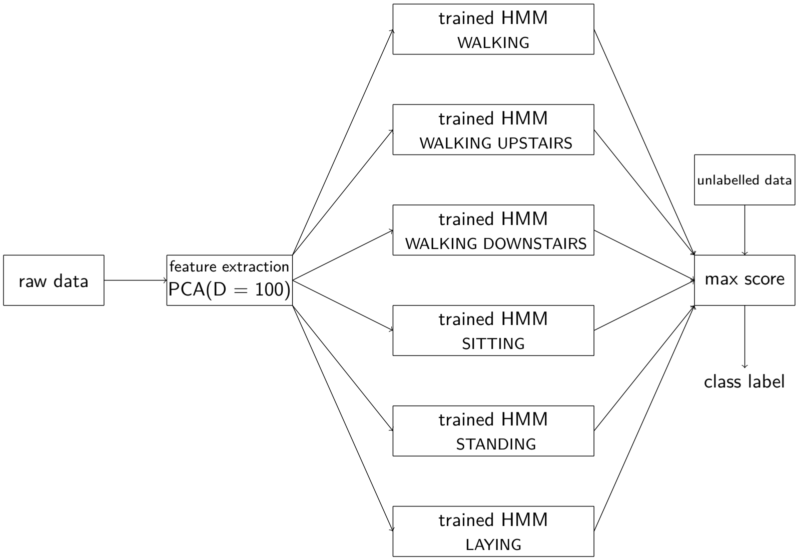


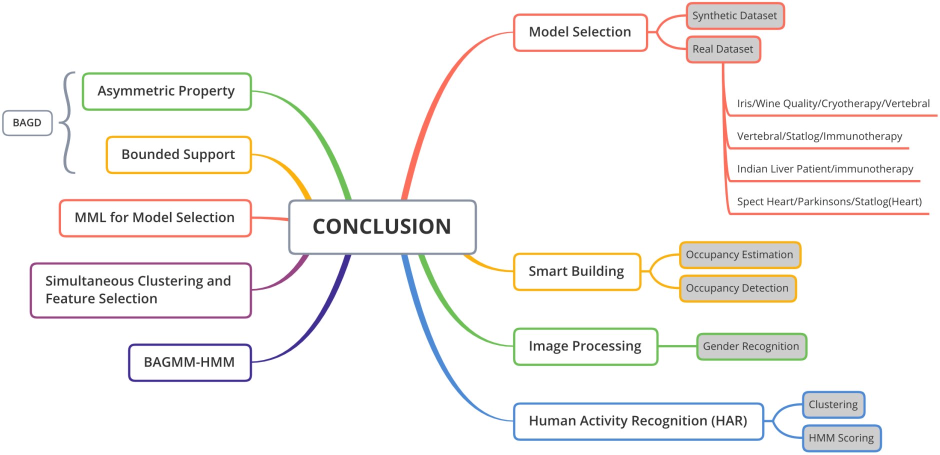
Figure: HMM for activity recognition according to the case of study

Table: Activity recognition results using different HMM models

	HMM Models			
Metrics	BAGMM-HMM	AGMM-HMM	BGMM-HMM	GMM-HMM
Accuracy	84.62%	77.27%	76.92%	75.00%
Precision	92.31%	69.32%	70.94%	69.44%
Recall	84.62%	77.27%	76.92%	75.00%
Specificity	97.20%	95.24%	24.57%	95.00%
F1-score	83.44%	71.21%	71.64%	68.88%
G-mean 1	88.38%	73.19%	73.87%	72.16%
G-mean 2	90.69%	85.79%	85.45%	84.40%
MCC	83.93%	69.98%	70.16%	68.02%

Our proposed model outperforms other HMMs, with the best configuration being $K = 2$ states and $M = 2$ mixture components associated with each state shown in Table 6, especially the highest accuracy of 84.64%.

Conclusion


CONCLUSION
BAGD
Asymmetric Property
Bounded Support
MML for Model Selection
Simultaneous Clustering and Feature Selection
BAGMM-HMM
Model Selection
Synthetic Dataset
Real Dataset
Iris/Wine Quality/Cryotherapy/Vertebral
Vertebral/Statlog/Immunotherapy
Indian Liver Patient/immunotherapy
Spect Heart/Parkinsons/Statlog(Heart)
Smart Building
Occupancy Estimation
Occupancy Detection
Image Processing
Gender Recognition
Human Activity Recognition (HAR)
Clustering
HMM Scoring

Publications I

Z. Xian, M. Azam, M. Amayri and N. Bouguila

Model Selection Criterion for Multivariate Bounded Asymmetric Gaussian Mixture Model

29*th* European Signal Processing Conference, EUSIPCO 2021

Z. Xian, M. Azam and N. Bouguila

Statistical Modeling Using Bounded Asymmetric Gaussian Mixtures: Application to Human Action and Gender Recognition

IEEE 22nd International Conference on Information Reuse and Integration for Data Science, *Best Student Paper Award*

Z. Xian, M. Azam, M. Amayri, W. Fan and N. Bouguila

Bounded Asymmetric Gaussian Mixture-Based Hidden Markov Models

Hidden Markov Models and Applications, Springer 2021

References I

- H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, pp. 716–723, December 1974.
- G. Schwarz *et al.*, "Estimating the dimension of a model," *The annals of statistics*, vol. 6, no. 2, pp. 461–464, 1978.
- H. Bozdogan, "Model selection and akaike's information criterion (aic): The general theory and its analytical extensions," *Psychometrika*, vol. 52, no. 3, pp. 345–370, 1987.
- J. Rissanen, *Stochastic complexity in statistical inquiry*, vol. 15. World scientific, 1998.

References II

- M. A. Figueiredo, J. M. Leitão, and A. K. Jain, "On fitting mixture models," in *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pp. 54–69, Springer, 1999.
- G. McLachlan and D. Peel, "Finite mixture models.,(john wiley & sons: New york.)," 2000.
- M. H. C. Law, M. A. T. Figueiredo, and A. K. Jain, "Simultaneous feature selection and clustering using mixture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 1154–1166, Sept 2004.

References III

T. Elguebaly and N. Bouguila, "Background subtraction using finite mixtures of asymmetric Gaussian distributions and shadow detection," *Machine Vision and Applications*, vol. 25, no. 5, pp. 1145–1162, 2014.

Q&A