

1 Experimental Settings

All experiments were run with the same hyperparameters for MAF, only varying the number of max training epochs for MAF between table 1 and table 2. For all experiments the MAF consisted of $blocks = 5$ MADE layers stacked, with $H = 512$ hidden neurons each, as this proved to be a good setting for both datasets used. A fairly small learning rate of $lr = 5 * 10^{-5}$ was chosen, with weight decay for the Adam optimizer being set to $wd = 1 * 10^{-6}$. The batch size was $b = 128$. Training data was shuffled to prevent batches consisting only of examples from one class. A patience of 5 was employed to stop training early if no improvement on the validation set in terms of log likelihood occurred within 5 epochs. If noise is *True*, Gaussian noise with $\mu = 0$ and $\sigma = 0.025$ was added directly to the raw data. Whereas, *augment* indicates that augmented samples, i.e. samples which were jittered with Gaussian noise ($\mu = 0, \sigma = 0.05$) and each variable being scaled with a factor drawn from a normal distribution ($\mu = 1, \sigma = 0.1$), were added to the training samples. Results are reported for the test set, which is the seventh subject of the Carrots data set, and the pre-made test set of UCIHAR respectively.

Dataset	Window	W_size	W_step	Augment	Noise	MVN LL	MAF LL	MVN ACC	MAF ACC
CARROTS	False	0	0	False	False	63.0739	68.2259	0.6006	0.5384
CARROTS	False	0	0	False	True	56.7594	59.6865	0.5932	0.6037
CARROTS	False	0	0	True	False	52.1412	68.1716	0.5683	0.5624
CARROTS	False	0	0	True	True	45.6814	59.1802	0.5103	0.6166
CARROTS	True	26	13	False	False	2986.2914	1555.2181	0.3299	0.3290
CARROTS	True	26	13	False	True	2016.7640	1416.1207	0.4398	0.3861
CARROTS	True	26	13	True	False	1881.0891	1720.2389	0.4312	0.4121
CARROTS	True	26	13	True	True	1566.8524	1513.4050	0.4909	0.5351
CARROTS	True	8	4	True	False	530.5136	695.9487	0.3759	0.5341
CARROTS	True	64	32	True	False	1375.9667	3319.6480	0.1215	0.4030
UCIHAR	True	128	64	False	False	717.9488	351.7945	0.9444	0.6291

Table 1: Log likelihoods and prediction accuracies of MAF and a baseline multivariate normal distribution trained per class with unconstrained covariance matrix. The Settings of the experiments are reported in the first 6 columns, the results in the remaining four. MAF was trained for a maximum of 2 epochs with the settings described above.

Dataset	Window	W_size	W_step	Augment	Noise	MVN LL	MAF LL	MVN ACC	MAF ACC
CARROTS	False	0	0	False	False	63.0739	47.1715	0.6006	0.3938
CARROTS	False	0	0	False	True	56.7594	61.4120	0.5932	0.6083
CARROTS	False	0	0	True	False	52.1412	53.1310	0.5683	0.3289
CARROTS	False	0	0	True	True	45.6814	60.8103	0.5103	0.6276
CARROTS	True	26	13	False	False	2986.2914	2860.8159	0.3299	0.4026
CARROTS	True	26	13	False	True	2016.7640	1951.0628	0.4398	0.3377
CARROTS	True	26	13	True	False	1881.0891	2965.4284	0.4312	0.4779
CARROTS	True	26	13	True	True	1566.8524	1972.0382	0.4909	0.3203
CARROTS	True	8	4	True	False	530.5136	1032.2761	0.3759	0.5125
CARROTS	True	64	32	True	False	1375.9667	5491.9022	0.1215	0.4051
UCIHAR	True	128	64	False	False	717.9488	749.2362	0.9444	0.9552

Table 2: Log likelihoods and prediction accuracies of MAF and a baseline multivariate normal distribution trained per class with unconstrained covariance matrix. The Settings of the experiments are reported in the first 6 columns, the results in the remaining four. MAF was trained for a maximum of 100 epochs with the settings described above.

Dataset	Window	W_size	W_step	Augment	Noise	MVN LL	MAF LL	MVN ACC	MAF ACC
MOSENSE	True	128	64	False	False	-242.2431	815.8154	0.4966	0.6375
MOSENSE	True	128	64	False	True	435.8816	718.1713	0.6580	0.7351
MOSENSE	True	128	64	True	False	587.4257	1080.5599	0.6731	0.7287
MOSENSE	True	128	64	True	True	210.0161	911.7357	0.6544	0.8089
MOSENSE	True	64	32	True	False	746.5845	721.6906	0.8182	0.7725
MOSENSE	True	32	16	True	False	371.3542	489.9848	0.7379	0.7393
UCIHAR	True	0	0	False	False	846.4878	913.8025	0.9474	0.8341
UCIHAR	True	0	0	False	True	846.4878	706.9777	0.9474	0.8456

Table 3: Log likelihoods and prediction accuracies of MAF and a baseline multivariate normal distribution trained per class with unconstrained covariance matrix. The Settings of the experiments are reported in the first 6 columns, the results in the remaining four. MAF was trained for a maximum of 10 epochs with the settings described above.

Dataset	Window	W_size	W_step	Augment	Noise	MVN LL	MAF LL	MVN ACC	MAF ACC
MOSENSE	True	128	64	False	False	-242.2431	1264.8010	0.4966	0.7487
MOSENSE	True	128	64	False	True	435.8816	1200.5637	0.6580	0.8254
MOSENSE	True	128	64	True	False	587.4257	1582.0021	0.6731	0.8030
MOSENSE	True	128	64	True	True	210.0161	1284.2458	0.6544	0.7916
MOSENSE	True	64	32	True	True	579.2101	868.5525	0.8166	0.8046
MOSENSE	True	32	16	True	True	289.0209	516.7429	0.7376	0.7864
UCIHAR	True	0	0	False	False	846.4878	1256.1664	0.9474	0.8792
UCIHAR	True	0	0	False	True	846.4878	868.8217	0.9474	0.9433

Table 4: Log likelihoods and prediction accuracies of MAF and a baseline multivariate normal distribution trained per class with unconstrained covariance matrix. The Settings of the experiments are reported in the first 6 columns, the results in the remaining four. MAF was trained for a maximum of 50 epochs with the settings described above.

Dataset	Window	W_size	W_step	Augment	Noise	MVN LL	MAF LL	MVN ACC	N
CARROTS	False	0	0	False	False	65.5414	47.1715	0.6006	
CARROTS	False	0	0	False	True	59.4133	61.4120	0.5932	
CARROTS	False	0	0	True	False	54.8067	53.1310	0.5683	
CARROTS	False	0	0	True	True	48.3399	60.8103	0.5103	
CARROTS	True	26	13	False	False	2988.2551	2860.8159	0.3299	
CARROTS	True	26	13	False	True	2018.9185	1951.0628	0.4398	
CARROTS	True	26	13	True	False	1883.5380	2965.4284	0.4312	
CARROTS	True	26	13	True	True	1569.2153	1972.0382	0.4909	
CARROTS	True	8	4	True	False	534.0041	1032.2761	0.3759	
CARROTS	True	64	32	True	False	1378.5929	5491.9022	0.1215	
MOSENSE	True	128	64	False	False	-240.8046	1485.4197	0.4966	
MOSENSE	True	128	64	False	True	437.3264	1316.1866	0.6580	
MOSENSE	True	128	64	True	False	588.9046	1621.9535	0.6731	
MOSENSE	True	128	64	True	True	211.4692	1316.2927	0.6544	
MOSENSE	True	64	32	False	True	874.2858	876.4729	0.8222	
MOSENSE	True	64	32	True	True	580.9881	894.1850	0.8166	
MOSENSE	True	32	16	False	True	442.3444	532.6112	0.7380	
MOSENSE	True	32	16	True	True	290.8765	541.4050	0.7376	
UCIHAR	True	128	64	False	False	-168.4130	1375.5543	0.8982	

Dataset	Window	W_size	W_step	Augment	Noise	MVN LL	MAF LL	MVN ACC	N
CARROTS	False	0	0	False	False	65.5414	68.2259	0.6006	
CARROTS	False	0	0	False	True	59.4133	59.6865	0.5932	
CARROTS	False	0	0	True	False	54.8067	68.1716	0.5683	
CARROTS	False	0	0	True	True	48.3399	59.1802	0.5103	
CARROTS	True	26	13	False	False	2988.2551	1555.2181	0.3299	
CARROTS	True	26	13	False	True	2018.9185	1416.1207	0.4398	
CARROTS	True	26	13	True	False	1883.5380	1720.2389	0.4312	
CARROTS	True	26	13	True	True	1569.2153	1513.4050	0.4909	
CARROTS	True	8	4	True	False	534.0041	695.9487	0.3759	
CARROTS	True	64	32	True	False	1378.5929	3319.6480	0.1215	
MOSENSE	True	128	64	False	False	-240.8046	21.6240	0.4966	
MOSENSE	True	128	64	False	True	437.3264	-25.8698	0.6580	
MOSENSE	True	128	64	True	False	588.9046	355.7859	0.6731	
MOSENSE	True	128	64	True	True	211.4692	294.3695	0.6544	
MOSENSE	True	64	32	False	True	874.2858	152.8148	0.8222	
MOSENSE	True	64	32	True	True	580.9881	275.8847	0.8166	
MOSENSE	True	32	16	False	True	442.3444	175.0598	0.7380	
MOSENSE	True	32	16	True	True	290.8765	245.8477	0.7376	
UCIHAR	True	128	64	False	False	-168.4130	419.3896	0.8982	