Feature Extraction with AutoEncoders

First Results

Presented by Gentry Atkinson

Data:

- MotionSense Dataset: Sensor Based Human Activity and Attribute Recognition
- Hosted on: https://www.kaggle.com/malekzadeh/motionsense-dataset/data
- Labeled iPhone accelerometer data from 24 subjects engaging in 6 different activities. 360 total files from 15 sec to 8 min at 50 Hz.
- Data has 12 dimensions:
  - Attitude on X, Y, Z axis
  - Gravity on X, Y, Z axis
  - Rotation on X, Y, Z axis
  - User Acceleration on X, Y, Z axis
- All information in Mobile Sensor Data Anonymization, (IoTDI ’19)

Data Processing:

- 9 of 12 features in the original set are derived features produced by proprietary Apple processes.
- The magnitude of the "User Acceleration" features is calculated with:
  - $\sqrt{x^2+y^2+z^2}$
- Samplest are trimmed to 370 values (the length of the shortest single file)

AutoEncoder Construction:

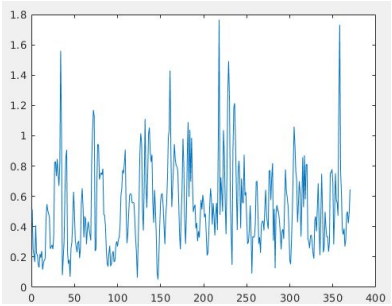
- Built using Keras w/ TensorFlow
- Encoder:
  - Input
  - Convolutional1D
  - Dense
- Decoder:
  - Dense
  - TransposeConvolutional1D
  - Output
- Trained over 10 epochs
- Decoder is removed so that now that hidden layer is the output of the network.

Feature Processing:

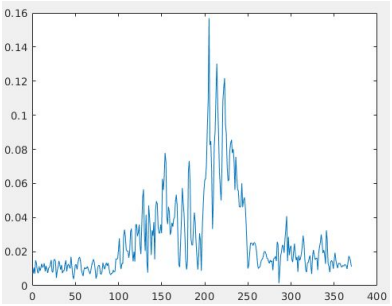
- A 128 value feature set is extracted for each 370 value signal.
- The features were clustered in Matlab using 6-means clustering.
- A table was produced with labels as rows and clusters as columns.
- A strong result is when there is one dominant cluster in every label.
- Several combinations of loss functions and optimizers were tried

Raw Data Visualizations:

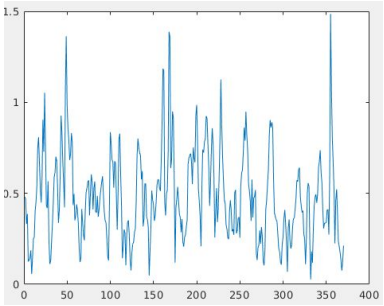
1. Traveling downstairs



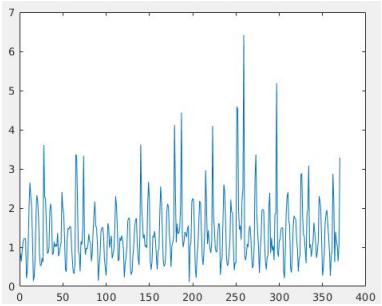
3. Sitting



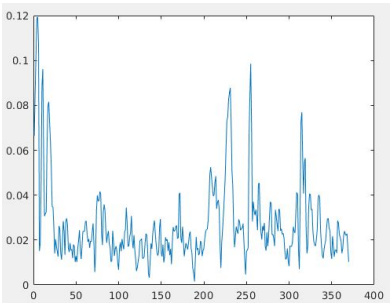
5. Traveling upstairs



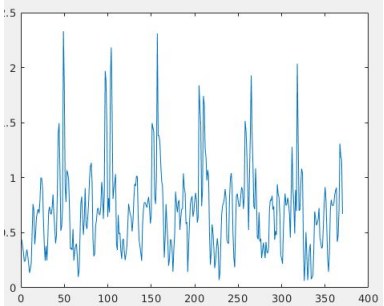
2. Jogging



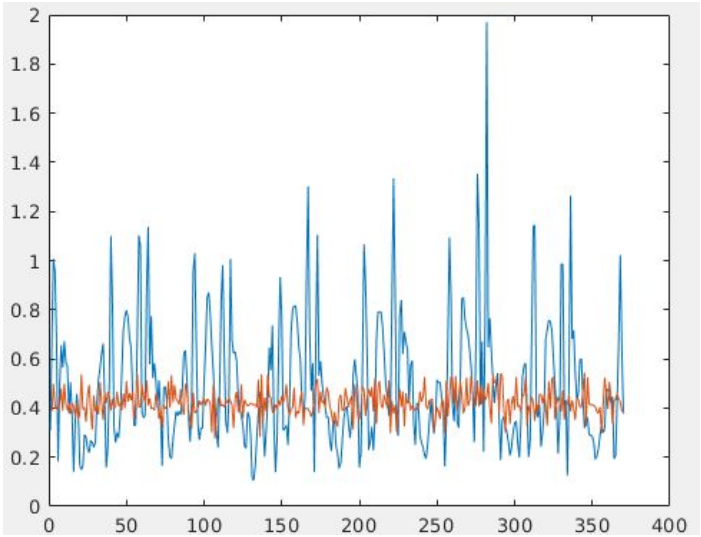
4. Standing



6. Walking



Mean Square Error:

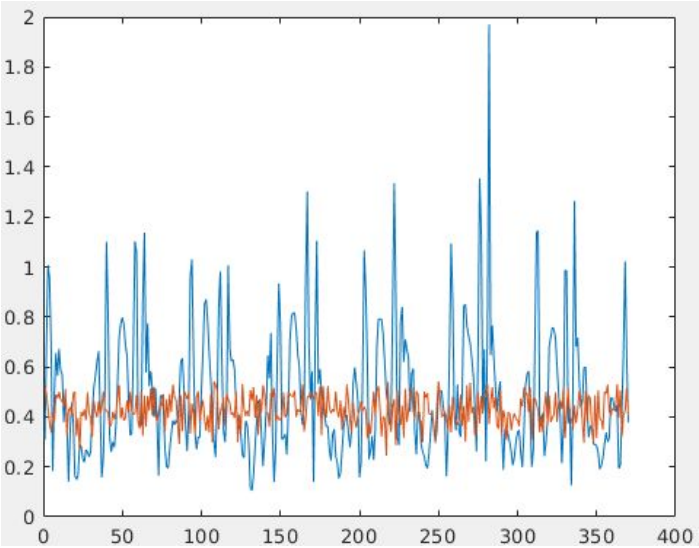



```
%Label 1: Downstairs, 58 samples
%Label 2: Jog, 38 samples
%Label 3: Sit, 39 samples
%Label 4: Stand, 38 samples
%Label 5: Upstairs, 58 samples
%Label 6: Walking, 57 samples
```

	1	2	3	4	5	6
1	14	0	0	13	27	4
2	0	0	0	38	0	0
3	0	39	0	0	0	0
4	0	38	0	0	0	0
5	15	0	10	7	8	18
6	6	0	7	13	21	10

	1	2	3	4	5	6
1	0.2414	0	0	0.2241	0.4655	0.0690
2	0	0	0	1	0	0
3	0	1	0	0	0	0
4	0	1	0	0	0	0
5	0.2586	0	0.1724	0.1207	0.1379	0.3103
6	0.1053	0	0.1228	0.2281	0.3684	0.1754

Huber Loss:

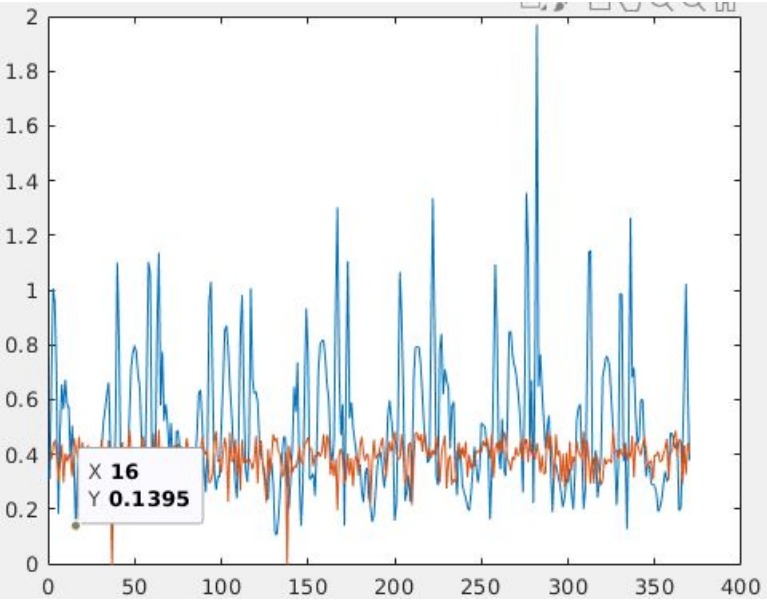



```
%Label 1: Downstairs, 58 samples
%Label 2: Jog, 38 samples
%Label 3: Sit, 39 samples
%Label 4: Stand, 38 samples
%Label 5: Upstairs, 58 samples
%Label 6: Walking, 57 samples
```

	1	2	3	4	5	6
1	0	29	5	15	9	0
2	0	0	36	2	0	0
3	39	0	0	0	0	0
4	38	0	0	0	0	0
5	0	9	1	7	29	12
6	1	12	1	24	12	7

	1	2	3	4	5	6
1	0	0.5000	0.0862	0.2586	0.1552	0
2	0	0	0.9474	0.0526	0	0
3	1	0	0	0	0	0
4	1	0	0	0	0	0
5	0	0.1552	0.0172	0.1207	0.5000	0.2069
6	0.0175	0.2105	0.0175	0.4211	0.2105	0.1228

Mean Square Logarithmic Error:

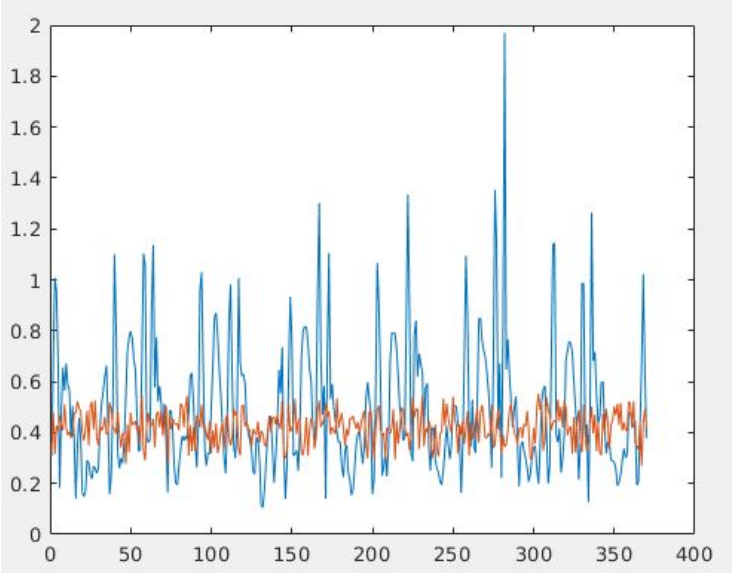



```
%Label 1: Downstairs, 58 samples
%Label 2: Jog, 38 samples
%Label 3: Sit, 39 samples
%Label 4: Stand, 38 samples
%Label 5: Upstairs, 58 samples
%Label 6: Walking, 57 samples
```

	1	2	3	4	5	6
1	12	0	0	13	24	9
2	38	0	0	0	0	0
3	0	0	39	0	0	0
4	0	0	38	0	0	0
5	5	12	0	8	6	27
6	7	7	1	21	9	12

	1	2	3	4	5	6
1	0.2069	0	0	0.2241	0.4138	0.1552
2	1	0	0	0	0	0
3	0	0	1	0	0	0
4	0	0	1	0	0	0
5	0.0862	0.2069	0	0.1379	0.1034	0.4655
6	0.1228	0.1228	0.0175	0.3684	0.1579	0.2105

MSE + Bias:

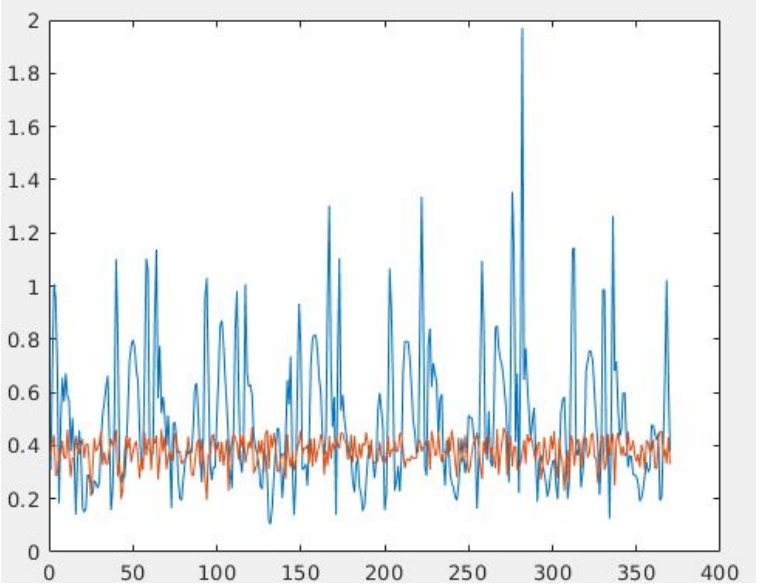



```
%Label 1: Downstairs, 58 samples
%Label 2: Jog, 38 samples
%Label 3: Sit, 39 samples
%Label 4: Stand, 38 samples
%Label 5: Upstairs, 58 samples
%Label 6: Walking, 57 samples
```

	1	2	3	4	5	6
1	13	0	15	0	4	26
2	38	0	0	0	0	0
3	0	39	0	0	0	0
4	0	38	0	0	0	0
5	7	0	15	12	16	8
6	12	2	6	6	9	22

	1	2	3	4	5	6
1	0.2241	0	0.2586	0	0.0690	0.4483
2	1	0	0	0	0	0
3	0	1	0	0	0	0
4	0	1	0	0	0	0
5	0.1207	0	0.2586	0.2069	0.2759	0.1379
6	0.2105	0.0351	0.1053	0.1053	0.1579	0.3860

MSE + Adam (gradient descent with momentum):

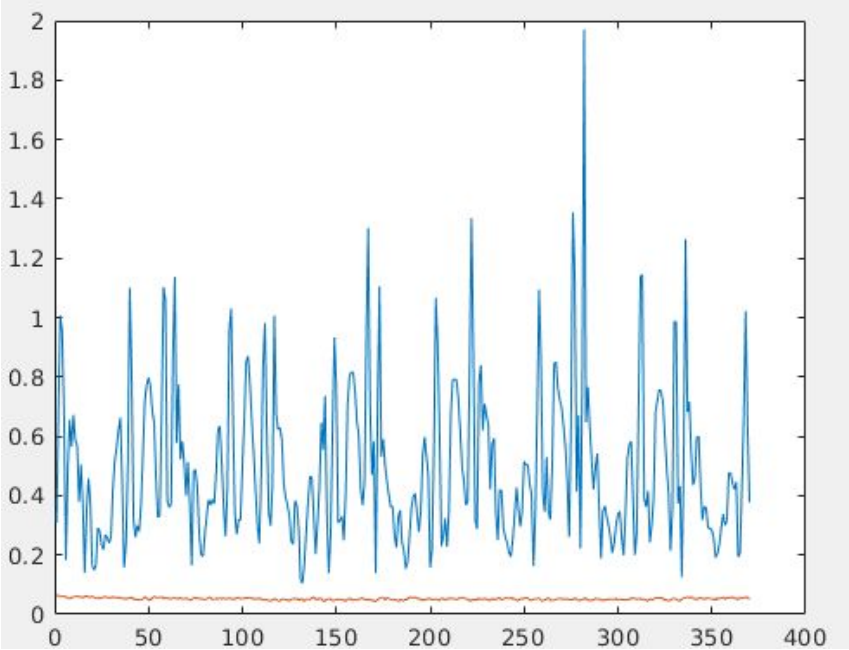



```
%Label 1: Downstairs, 58 samples
%Label 2: Jog, 38 samples
%Label 3: Sit, 39 samples
%Label 4: Stand, 38 samples
%Label 5: Upstairs, 58 samples
%Label 6: Walking, 57 samples
```

	1	2	3	4	5	6
1	5	0	16	13	0	24
2	0	0	38	0	0	0
3	0	39	0	0	0	0
4	0	38	0	0	0	0
5	14	0	7	18	10	9
6	10	0	23	6	6	12

	1	2	3	4	5	6
1	0.0862	0	0.2759	0.2241	0	0.4138
2	0	0	1	0	0	0
3	0	1	0	0	0	0
4	0	1	0	0	0	0
5	0.2414	0	0.1207	0.3103	0.1724	0.1552
6	0.1754	0	0.4035	0.1053	0.1053	0.2105

Cosine Proximity Loss:



Conclusions:

- The feature sets do show a preference to cluster in discernable groups.
- Sitting and Standing will be very difficult to differentiate with the current setup.
- The current model "flattens" data badly but some parameter combinations are doing a good job of preserving the "structure" of the raw data in the decoded signal.