**Machine Learning for Energy Segmentation and Forecasting**

Dominique Smith, REU Student, Arizona State University
Graduate Mentor: Kristen Jaskie, Faculty Advisor: Andreas Spanias
SenSIP Center, School of ECEE, Arizona State University

**PROBLEM STATEMENT**
- Energy usage can change over time
- People can move in or out of an area
- Useful to segment customers based off energy usage
- Can segments be used to improve energy forecasting?
- Compression helps reduce size of data, making it easier to store

**MOTIVATION**
- Smart meters produce large amounts of data
- Compression helps smooth data and minimize outliers
- Creating customer segments helps track and analyze energy usage over time
- Forecasting energy usage can help develop efficiency programs, plan energy grid development, consumer targeting, etc.

**EXPERIMENTAL METHODS**
- Considered raw data, compressed data, clustered data, and every combination
- Fed combinations into a predictive long short-term memory (LSTM) neural network
- Compared forecasts to measure effectiveness of clustering and compression

**RESULTS**
- Initially found six inherent clusters in the data
- Training the LSTM on the entire data set gives good results
- Compressing the data improves forecasting results by reducing the unpredictable randomness
- Clustering did not improve forecasting, but still useful for other segmentation tasks

**REFERENCES**


**Heatmap of energy use per customer segment [1]**

**Comparison of LSTM predictions to actual data [1]**

<table>
<thead>
<tr>
<th>MSE</th>
<th>Percent improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncompressed unclustered</td>
<td>0.0118</td>
</tr>
<tr>
<td>Compressed unclustered</td>
<td>0.0106</td>
</tr>
<tr>
<td>Uncompressed clustered</td>
<td>0.0165</td>
</tr>
<tr>
<td>Compressed clustered</td>
<td>0.0145</td>
</tr>
</tbody>
</table>

Mean squared error of predictions [1]