Applied Machine Learning

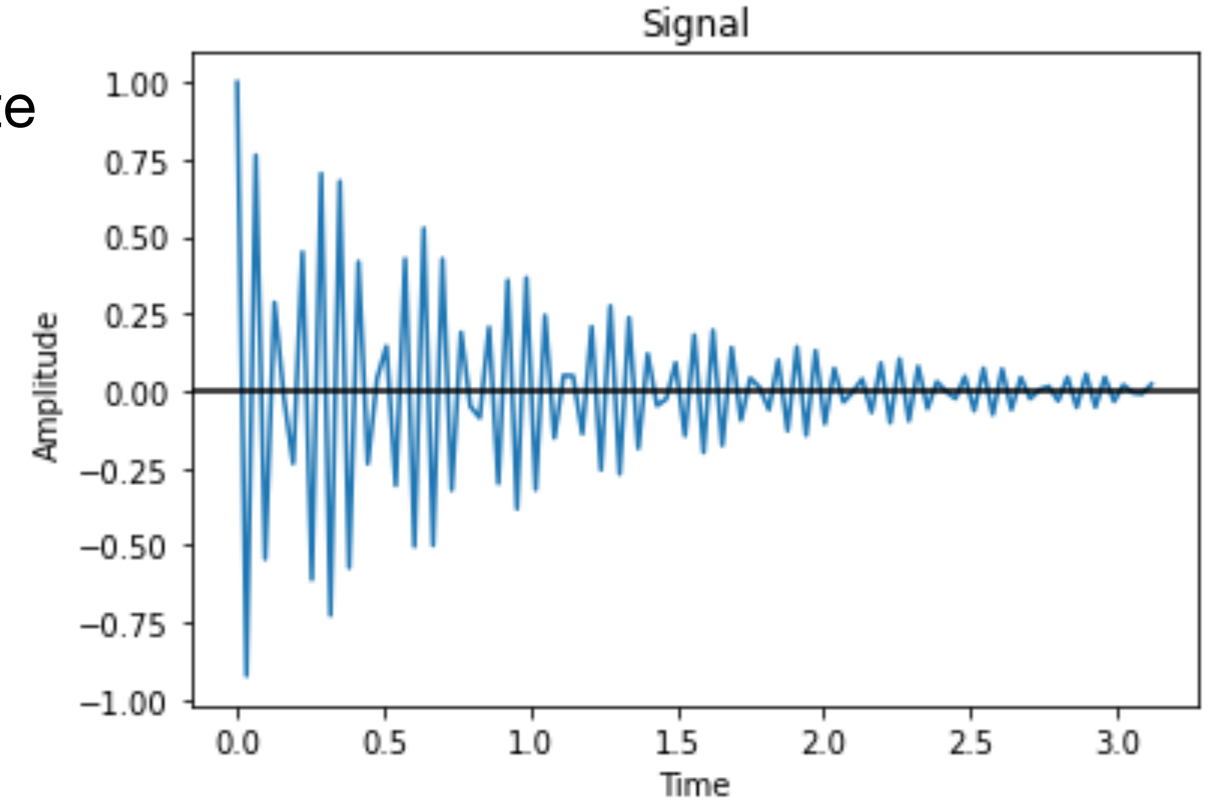
Vector Quantization

Vector Quantization

- Signals, datasets, and local patterns
- Codebooks and codewords
- Codebook construction

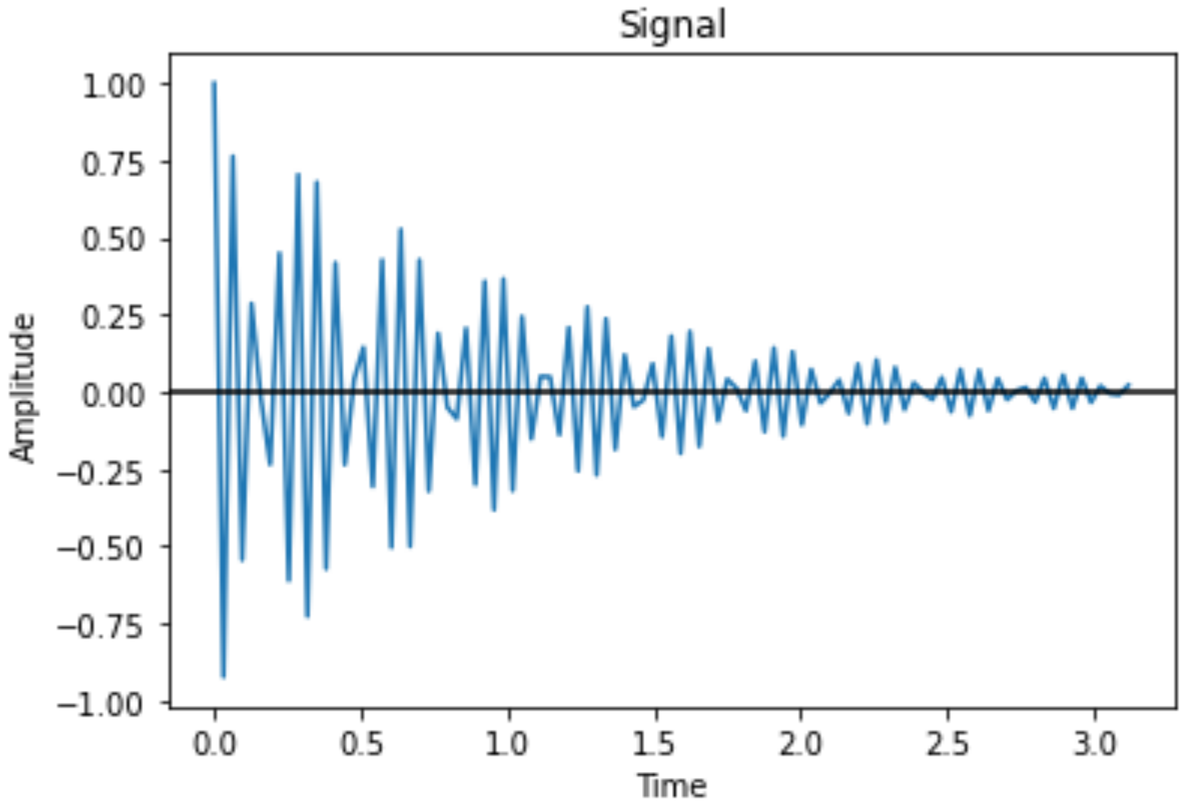
Vector Quantization - VQ

- Capture local patterns in signals
- Map range of values into a discrete set
- identification of local patterns in datasets
  - human senses
  - written text
  - textures in images



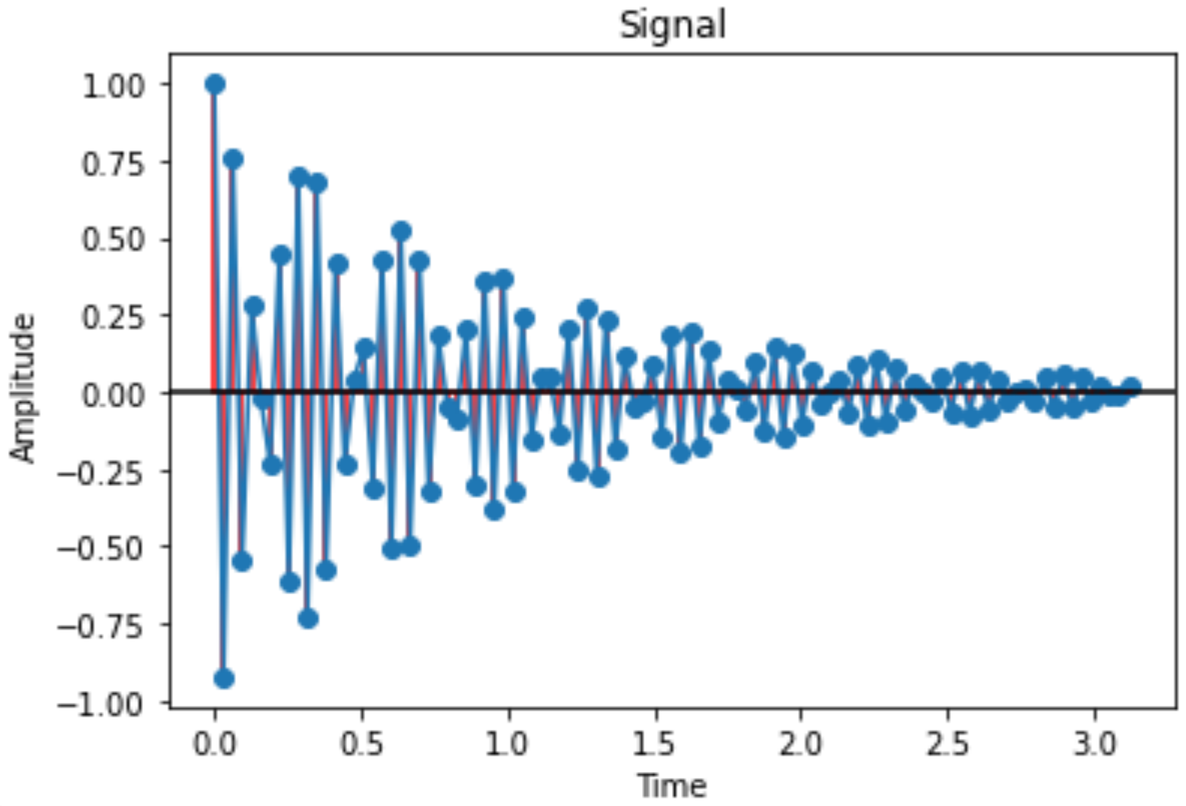
Vector Quantization - VQ

- Capture local patterns in signals
- Map range of values into a discrete set
- Speech recognition
  - phonemes
  - frequency components
- Inertial Measurement Units



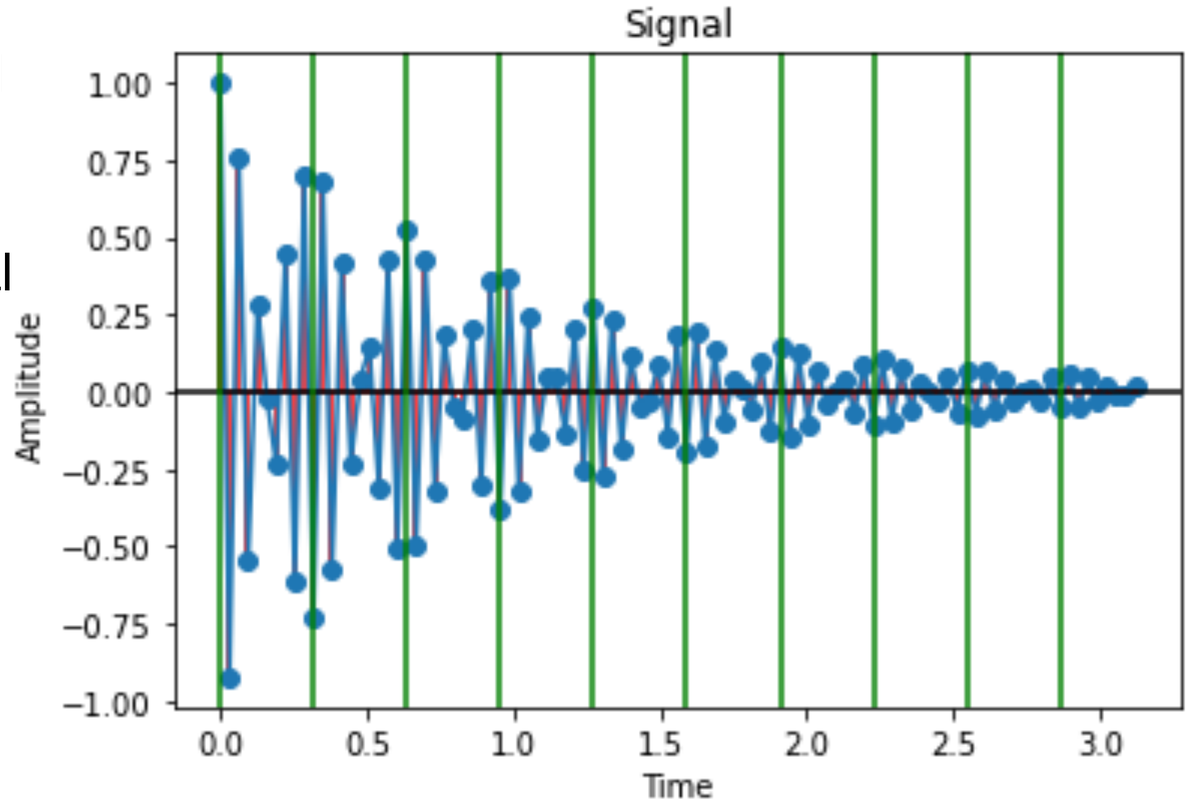
VQ Codebook and Codewords

- Local patterns may come in different shapes and sizes
- Repetitive
  - Build dictionary or codebook made of codewords
  - Feature vectors from identified codewords
- Composition based on codewords



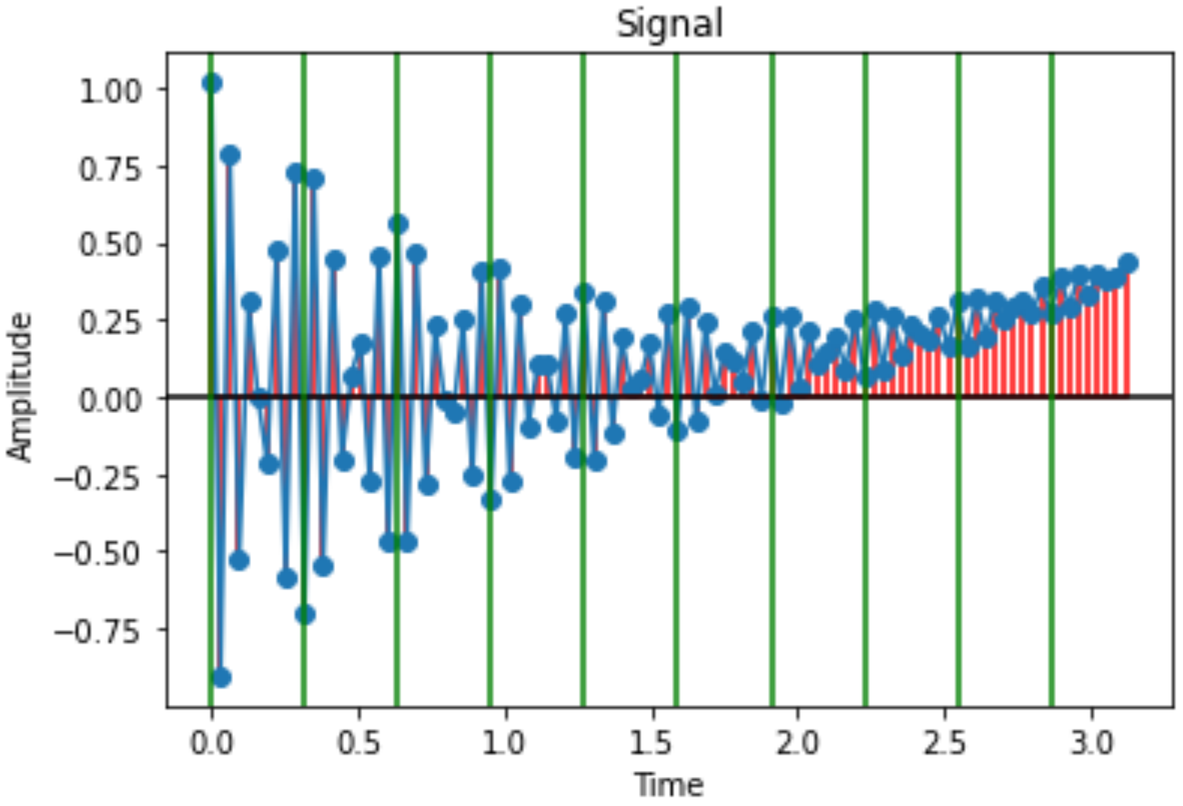
Construction of Codebook

- Codebook construction
  - Split each sequence into equally-sized patches
  - Convert each patch into d-dimensional vector
    - overlap allowed
    - each patch is an item
  - Codebook: clusters from dataset
    - One codeword per cluster: $[1 \ldots k]$



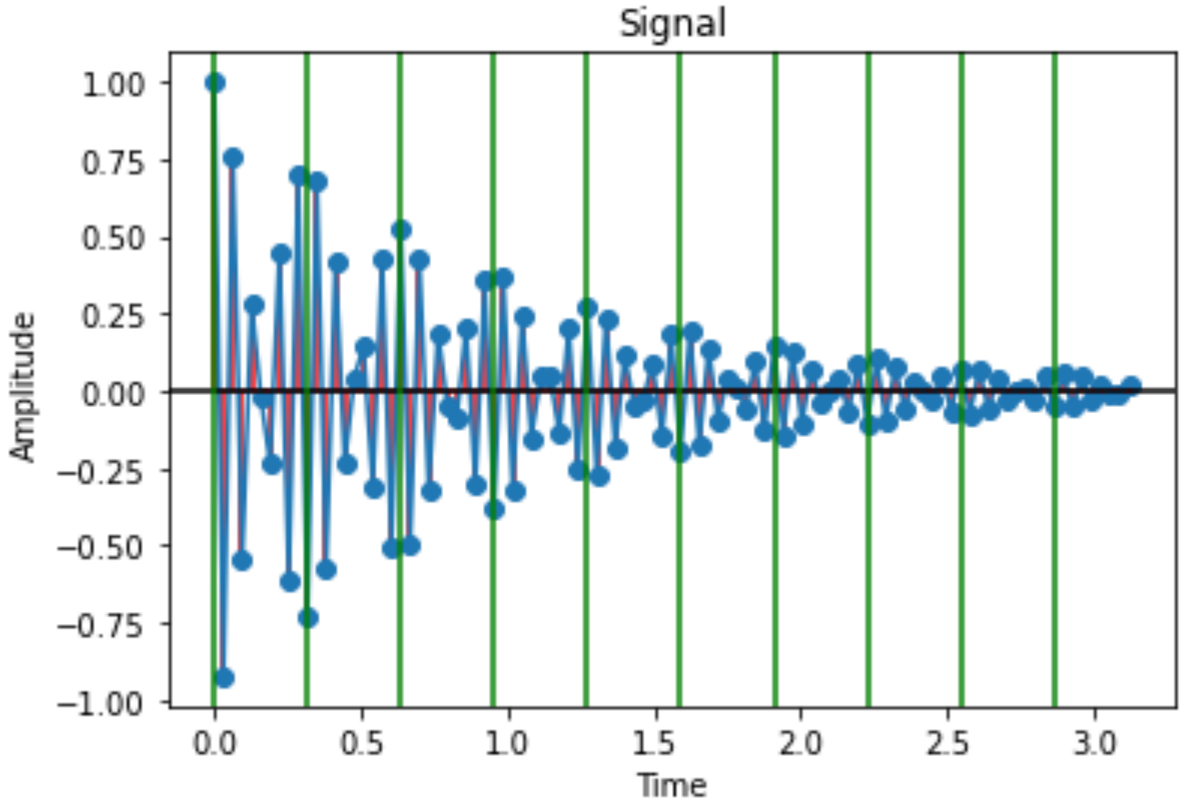
Characterization of new patch

- To codify new sequence:
  - Split new sequence into equally-sized patches
  - Convert each patch into d-dimensional vector
  - Associate vector to closest codeword
  - Build histogram of the k codewords for all the vectors
  - histogram with k features represents sequence



Patches in Different Domains

- Patches for d-dimensional vectors
  - Sound signals: d
  - Grayscale image patches: $\sqrt{d} \times \sqrt{d}$
  - Color image patches: $\sqrt{d/3} \times \sqrt{d/3} \times 3$
  - 3-axis accelerometer signals: $3 \times d/3$



Vector Quantization

- Signals, datasets, and local patterns
- Codebooks and codewords
- Codebook construction

Applied Machine Learning

Vector Quantization