Clustering Case Study: Forest Cover Types

Forest Cover Types

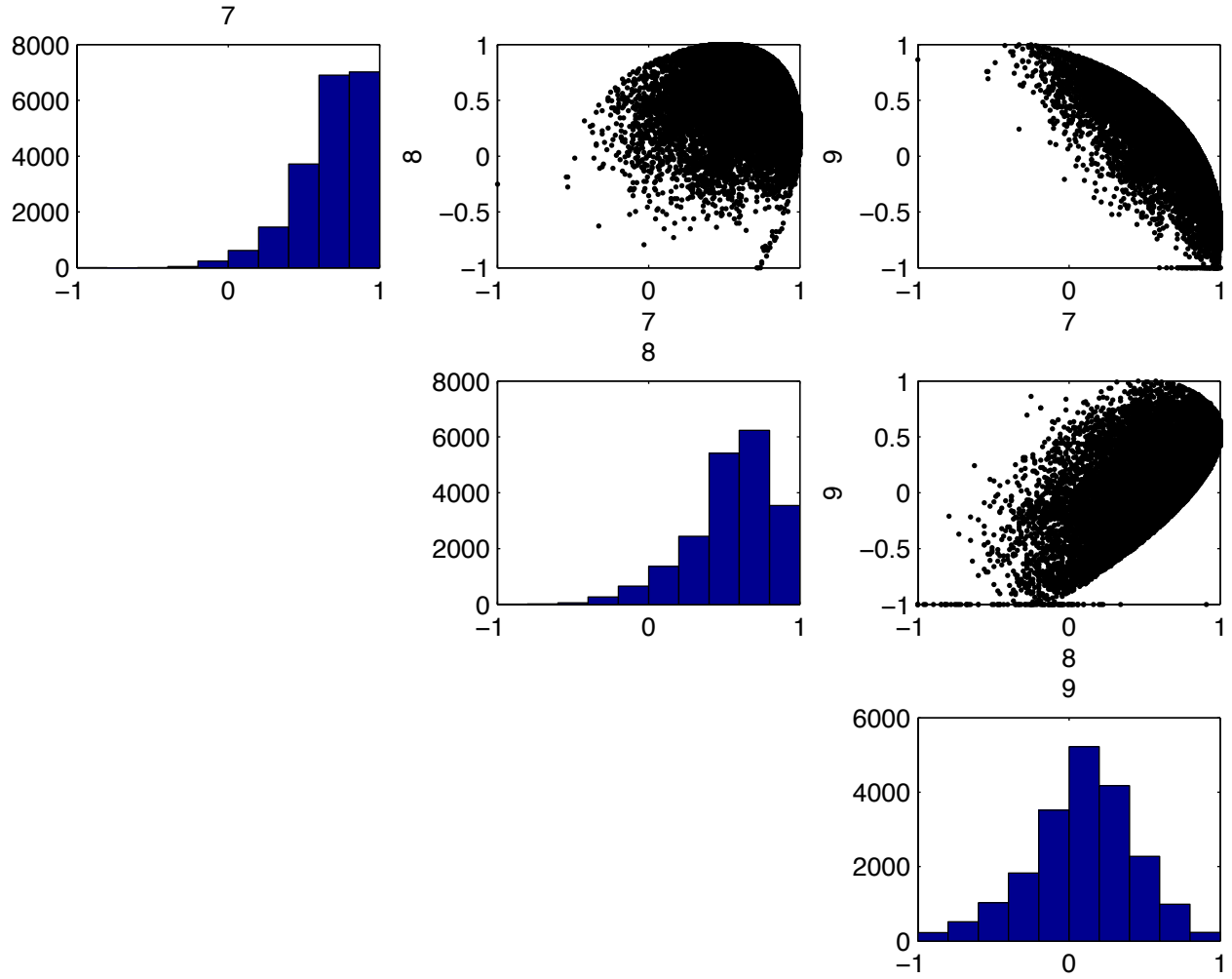
Label	Name
0	Krummholz
1	Spruce/Fir
2	Lodgepole Pine
3	Ponderosa Pine
4	Cottonwood/Willow
5	Aspen
6	Douglas-fir

26

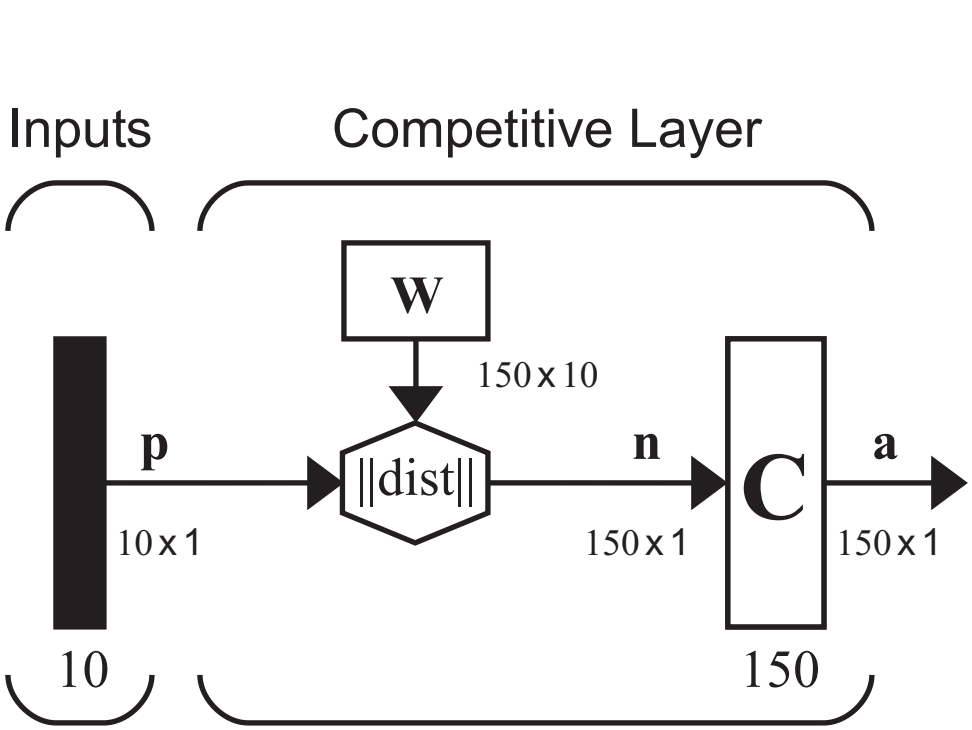
Independent Variables

Variable Number	Description	Units
1	Elevation in meters	meters
2	Aspect in degrees azimuth	azimuth
3	Slope in degrees	degrees
4	Horz Dist to nearest surface water	meters
5	Vert Dist to nearest surface water	meters
6	Horz Dist to nearest roadway	meters
7	Hillshade index at 9am, summer solstice	0 to 255 index
8	Hillshade index at noon, summer solstice	0 to 255 index
9	Hillshade index at 3pm, summer solstice	0 to 255 index
10	Horz Dist to nearest wildfire ignition points	meters

Scatter Plots (Variables 7, 8, 9)

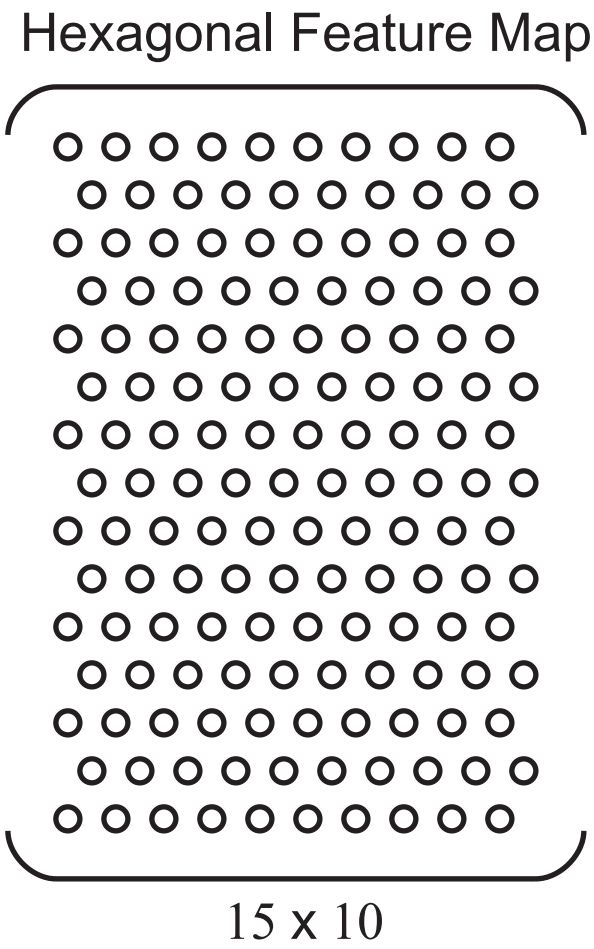


SOM Network Architecture



$n_i = \|_i\mathbf{w}\text{-}\mathbf{p}\|$

$\mathbf{a} = \mathbf{compet}(\mathbf{n})$



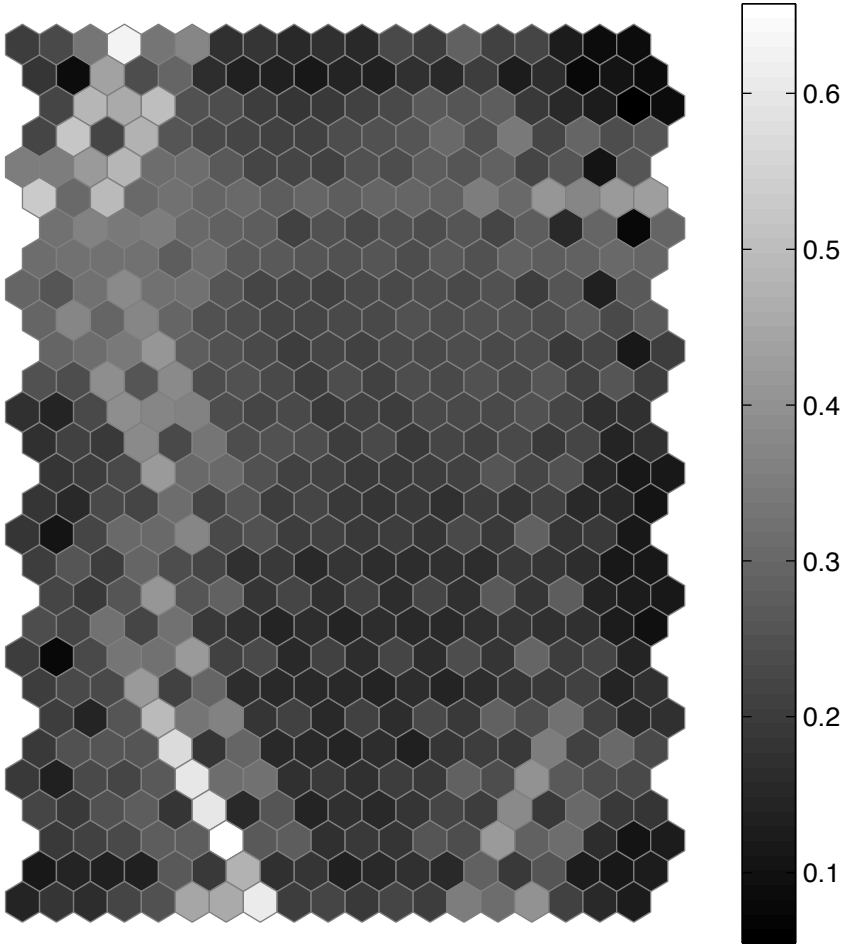
Batch SOM Algorithm

$$ {}_i\mathbf{w}(q) = {}_i\mathbf{w}(q-1) + h_{i^*,i}\left(\mathbf{p}(q) - {}_i\mathbf{w}(q-1)\right) $$

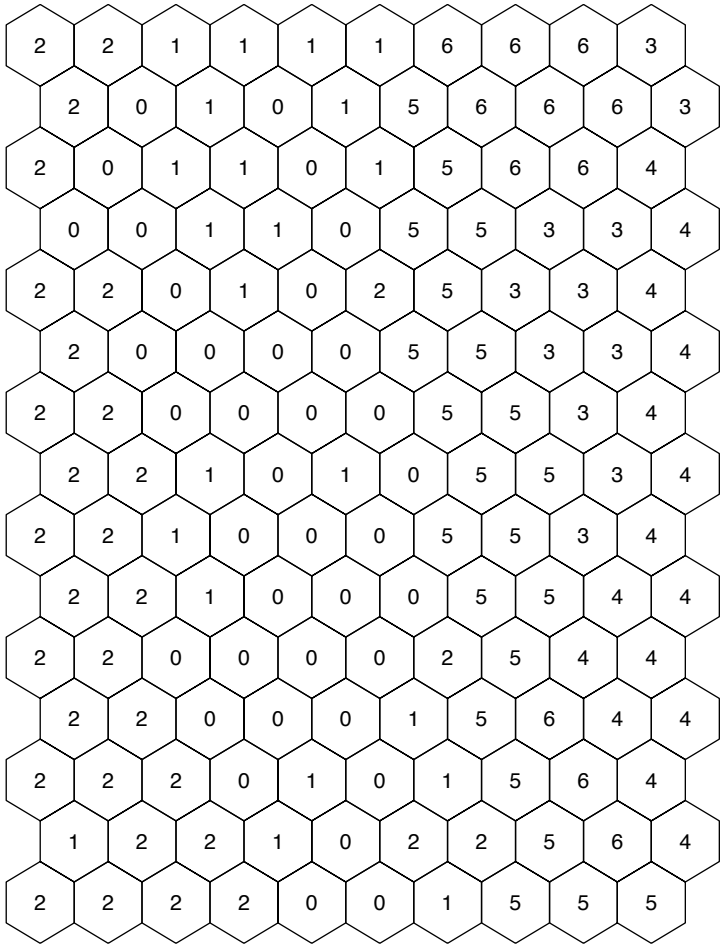
$$ h_{i^*,i} = \begin{cases} \alpha & i \in N_{i^*}(d) \\ 0 & i \notin N_{i^*}(d) \end{cases} $$

$$ {}_i\mathbf{w}(k) = \frac{\sum_{q=1}^{Q} h_{i^*(q),i}\mathbf{p}(q)}{\sum_{q=1}^{Q} h_{i^*(q),i}} $$

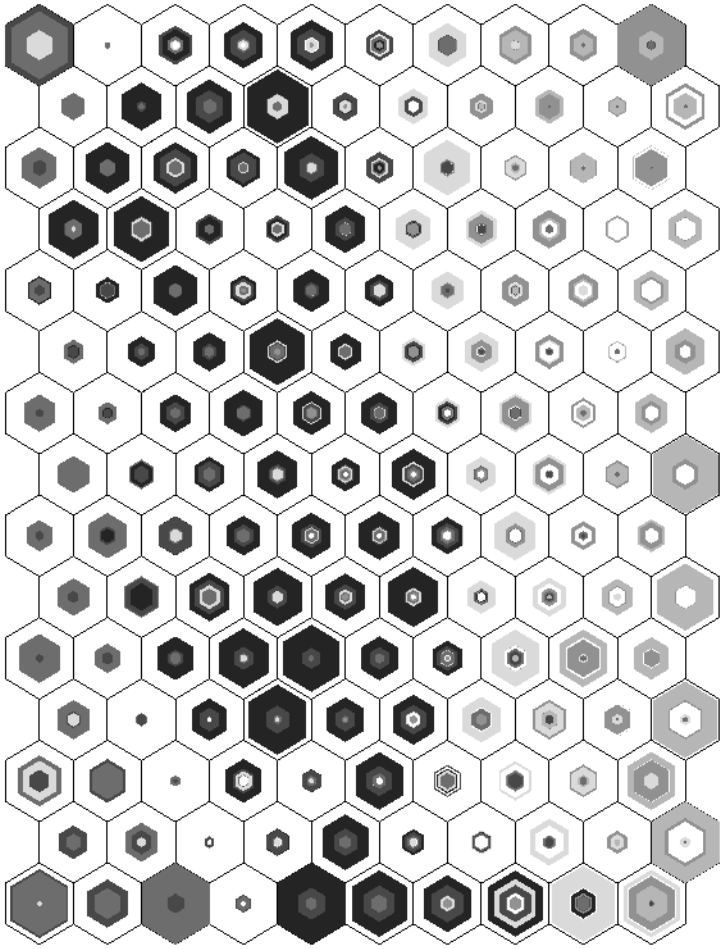
U-Matrix (Neuron Distances)



Labeled SOM


2 2 1 1 1 1 6 6 6 3
2 0 1 0 1 5 6 6 6 3
2 0 1 1 0 1 5 6 6 4
0 0 1 1 0 5 5 3 3 4
2 2 0 1 0 2 5 3 3 4
2 0 0 0 0 5 5 3 3 4
2 2 0 0 0 0 5 5 3 4
2 2 1 0 1 0 5 5 3 4
2 2 1 0 0 0 5 5 3 4
2 2 1 0 0 0 5 5 4 4
2 2 0 0 0 0 2 5 4 4
2 2 0 0 0 1 5 6 4 4
2 2 2 0 1 0 1 5 6 4
1 2 2 1 0 2 2 5 6 4
2 2 2 2 0 0 1 5 5 5

Hit Histogram



Darkest – Type 0
Lightest – Type 6

Component Planes