Appendix: Uni-perceptron (augmented bi-perceptron, see paper C.42)
A Parallel Bi-perceptron Approach and its Application to Data Classification,
by Daw-Ran Liou, Yang-En Chen, Cheng-Yuan Liou, 2016

This article is an appendix of the bi-perceptron paper. It shows how to operate the bi-perceptron technique with half strips to accomplish the classification tasks.

.

Uni-perceptron (augmented bi-perceptron, see paper C.42)
Case 1: for discretized binary patterns.

Data set MNIST; $\{p_n , n=1 \sim 60000=N\}$;
Number of dimensions (attributes)=D=28*28 pixels=784 pixels (B/W).
https://scidm.nchc.org.tw/dataset/mnist (handwritten digits; 0,1,..,9)
There are total 10 classes in the data set, $\{C_0, C_1, C_2, .., C_j, \dots , C_{J=9}\}$

We plan to construct multiple perceptrons to discriminate one digit '0' from all other digits, '1',..,'9'. The whole procedure is in the following context. We iteratively operate the following 5 Steps for the class of digit '0'

Step 1.
Construct two sets, $\{p_n^1\}$ and $\{ p_n^2\}$, for each. pattern, $p_n \in C_j$. Here we set $C_j = C_0$.
Set $\{p_n^1\}$ contains 784 patterns. These 784 patterns are in the classes that are different from C_j and they are the nearest neighboring patterns to p_n.
Set $\{p_n^1\}=\{p_m ; p_m \in \{\min |p_n - p_m|, m=1\sim784\}$ and $p_m \notin C_j\}$.
Construct an ancillary hyperplane (784 dimensions), W_a, that passes these784 patterns in $\{p_n^1\}$. Let p_n on the positive side of this ancillary hyperplane. Construct a normal vector that passes the pattern p_n and is perpendicular to this ancillary hyperplane W_a. Let p_n be the end point of this normal vector.

Step 2.
Construct a set, $\{p_n^2\}$, that contains all patterns that belong to the same class C_j and are on the same positive side of this ancillary hyperplane as p_n. Suppose the number of patterns in the set $\{p_n^2\}$ is $|\{p_n^2\}|$. So, there are total number, $|C_j|$, of such sets as $\{p_n^1\}$ for the class C_j and total number,

$|C_j|$, of sets $\{p_n{}^2\}$.

Step 3.
For the class C_j, pick the set that has the largest amount of patterns, $|\{p_n{}^2\}^1| = \max\{|\{p_n{}^2\}|; p_n \in C_j\}$. We will construct the first perceptron for the class C_j by using the patterns in $\{ p_n{}^2\}^1$.

Step 4.
Find the pattern, p_u, in $\{ p_n{}^2\}^1$ that is the nearest pattern to its ancillary hyperplane, $p_u \in \{p_n{}^2\}^1$. Construct a discrimination hyperplane, W, that passes the middle point of the normal vector and parallels the ancillary hyperplane. Note that the end point of the normal vector is the pattern p_u. The hyperplane, W^1, is right on the middle point of the normal vector. The weights of a neuron can be obtained from the hyperplane W.

Step 5, Delete all patterns $\{p_n{}^2\}^1$ from the class C_j and construct a reduced set C_j^1 for the class C_j. Then, use the reduced set C_j^1 in the five steps from Step 1 to Step 5 to construct the next discrimination perceptron W^2 from $\{p_n{}^2\}^2$.

Iteratively operate the five steps to further reduce the set C_j, $\{C_j^1, C_j^2, C_j^3,\ldots, C_j^T\}$, and obtain discrimination perceptrons, W^1, W^2, W^3,.., W^T, where T denotes the last iteration.
Note that $\{p_n{}^2\}^a \cap \{p_n{}^2\}^b = \varnothing$ for $a \neq b$, and $|\{p_n{}^2\}^1| \geq |\{p_n{}^2\}^2| \geq |\{p_n{}^2\}^3| \geq \ldots |\{p_n{}^2\}^T|$. The union of all reduced sets contain all patterns in the class C_j, $\{\{p_n{}^2\}^1 \cup \{p_n{}^2\}^2 \cup \{p_n{}^2\}^3 \cup \ldots \cup \{p_n{}^2\}^T\} = C_j$. This procedure is similar to that for the bi-perceptron in paper C.42.

Case 2: for analog patterns.

When the pattern is not in discretized form, one can use the same technique as that of bi-perceptron with half strips.

The procedure starts from the construction of an ancillary hyperplane, W_a, for the outmost pattern, $p_{outmost}$, which is the most distant pattern from the average center of all patterns. Let $p_{outmost} \in C_j$. Let W_a be perpendicular to the line between the pattern, $p_{outmost}$, and the center. Note W_a passes the pattern $p_{outmost}$. Then parallelly shift this hyperplane, W_a, to its limit to the patterns in the set $\{p_{outmost}{}^{1}\}^{1}$, that are in the different classes from the outmost pattern class, C_j, and are the nearest patterns to W_a. All patterns on the same side of this ancillary hyperplane as the outmost patter will be included in a set, $\{p_{outmost}{}^{2}\}^{1} = \{p; p \in$ on the same side of W_a as $p_{outmost}\}$. Note that all patterns in the set $\{p_{outmost}{}^{2}\}^{1}$ belong to the class C_j. Then find a pattern from this set, $p_u \in \{p_{outmost}{}^{2}\}^{1}$, that is the nearest pattern to the ancillary perceptron W_a. Construct a hyperplane, W, for the set $\{p_{outmost}{}^{2}\}^{1}$ that is right on the middle point between this nearest pattern, p_u, and W_a. Note that W is parallel to W_a.

We then delete all patterns in $\{p_{outmost}{}^{2}\}^{1}$ from C_j and obtain a reduced set C_j^1, $C_j=C_j^1 \cup \{p_{outmost}{}^{2}\}^{1}$. Apply the same procedure for the reduced set C_j^1, one can construct the next perceptron for the outmost pattern in the reduced set C_j^1.

We iteratively operate this procedure to further reduce the set C_j, C_j^1 , C_j^2 , C_j^2 , C_j^3 …., C_j^T. This procedure stops with an empty set $C_j^{T+1} =\emptyset$.
All constructed perceptrons will be used in the first hidden layer in the network.
The outputs of the first hidden layer are all discretized binary digits with faithful representations (homogeneous representations in Chapter 4). One can apply the same technique as that for the discretized binary patterns for the binary outputs of the first hidden layer to construct the succeeding layer. The positive direction of the normal vector of each discriminative perceptron, W, in the first hidden layer is flexible.