Defence Research and Development Canada
Recherche et développement pour la défense Canada

Weight Adjustment Using Machine Learning Applied to The Analytical Hierarchy Process

(Caelum Kamps, Rahim Jassemi-Zargani)

International Symposium on The Analytical Hierarchy Process 2018
12th – 15th July 2018

Dr. Rahim Jassemi-Zargani
Space & ISR Applications (SIA) Section
DRDC Ottawa Research Centre
Ottawa, Canada

DRDC | RDDC

Canada

Background

Multi Criteria Decision Making (MCDM) for classification in **dynamic environments**

- Medical diagnosis
- Economic markets
- Self Driving Cars
- Various military applications

Dynamic environments require adaptability to **evolving and updating inputs** (cues).

When only some of the cues are available, the original **comparison matrices** can be truncated to generate new weights for **information deficient decision structures**.

Machine learning is used to adjust the weights of the information deficient decision structures to **improve classification accuracy**

Objective

The objective of this work is to address two areas in which the AHP can be supplemented to improve its reliability under specific conditions.

Those areas are:

1. **Information Deficiency**
2. **Interrelationships and Dependencies within the Cues**

The proposed solution combines the AHP with machine learning to produce improved classification accuracy of a dataset under information deficient conditions.

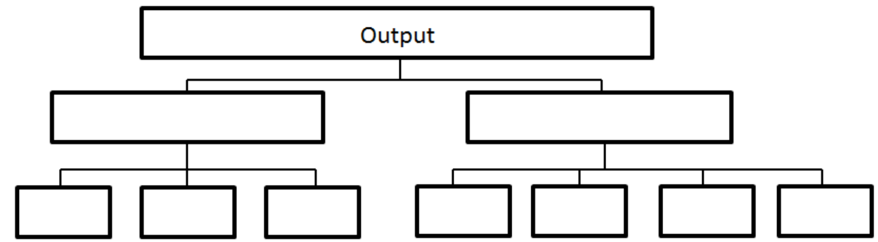
Information Deficiency

The graphics depict the change that happens to the decision structures under **deficient information** states

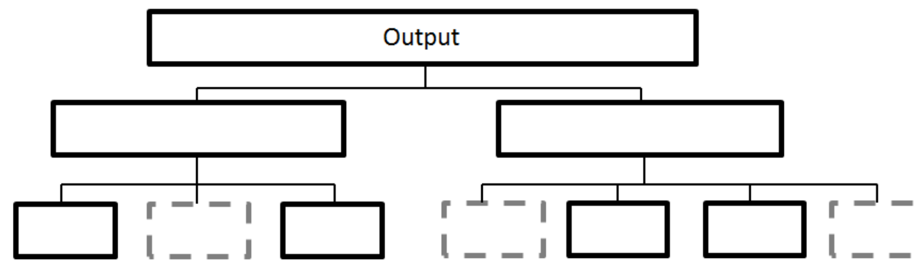
The weights of the **complete information** decision structure are generated through the **AHP**

The **comparison matrices** from the complete information state are **truncated** to produce initial weights for the **deficient information** state

Complete Information Decision Structure



Deficient Information Decision Structure



Cue Dependency and Correlations

The ability to calculate or measure certain cues can rely on the availability of others. These are known as **dependent cues**.

Under states of **information deficiency,** there usually exist **correlations** to be exploited between dependent cues and the cues they rely on. These relationships can also exist between **independent** cues.

These **correlations and relationships** can be exploited through the application of **machine learning** to the weights of **deficient information states**.

Self Driving Car Example

Independent cues:

- Range
- Direction
- Heading

Dependent cues:

- Angle of Approach (AOA)
- Closest Point of Approach (CPA)

Information State Space

All of the cues under consideration are labeled either **dependent or independent.**

Dependent cues are only available when their **underlying dependencies** are available.

The information state space is the list of all possible states of **availability** for the cues.

The table to the right represents a **complete information state space** for the self driving car example.

Self Driving Car Example

Independent cues:

- Range
- Direction
- Heading

Dependent cues:

- Angle of Approach (AOA)
- Closest Point of Approach (CPA)

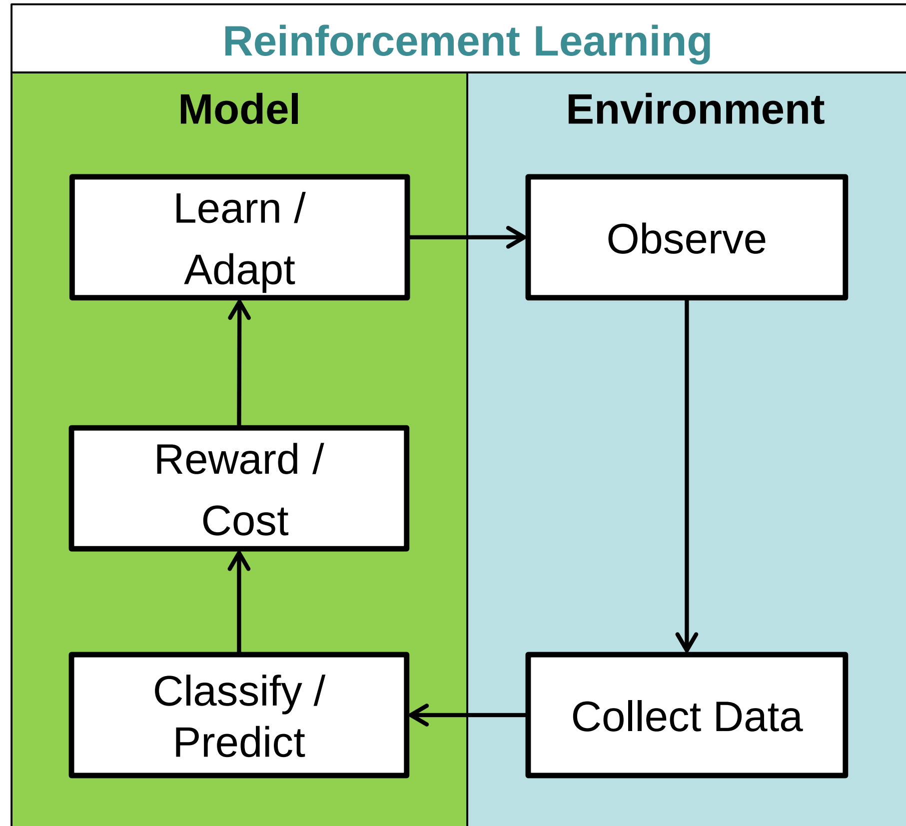
	Cue Availability					
Information State	Independent Cues			Dependent Cues		Number of Available Cues
	Range	Direction	Heading	AOA	CPA	
1	0	0	1	0	0	1
2	0	1	0	0	0	1
3	0	1	1	1	0	3
4	1	0	0	0	0	1
5	1	0	1	0	0	2
6	1	1	0	0	0	2
7	1	1	1	1	1	5

Machine Learning

Machine learning is a broad term. The specific type used in this work is **reinforcement learning**.

Reinforcement learning attempts to learn through **trail and error** by adjusting the model parameters based on a series of predictions followed by a **cost or reward** for that **prediction.**

This work uses **stochastic gradient descent,** an algorithm that minimizes a cost function over a series of predictions.

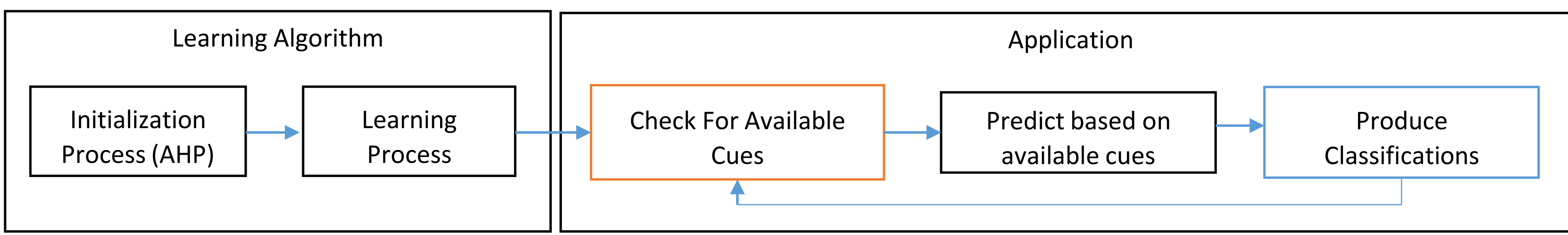


Algorithm Structure

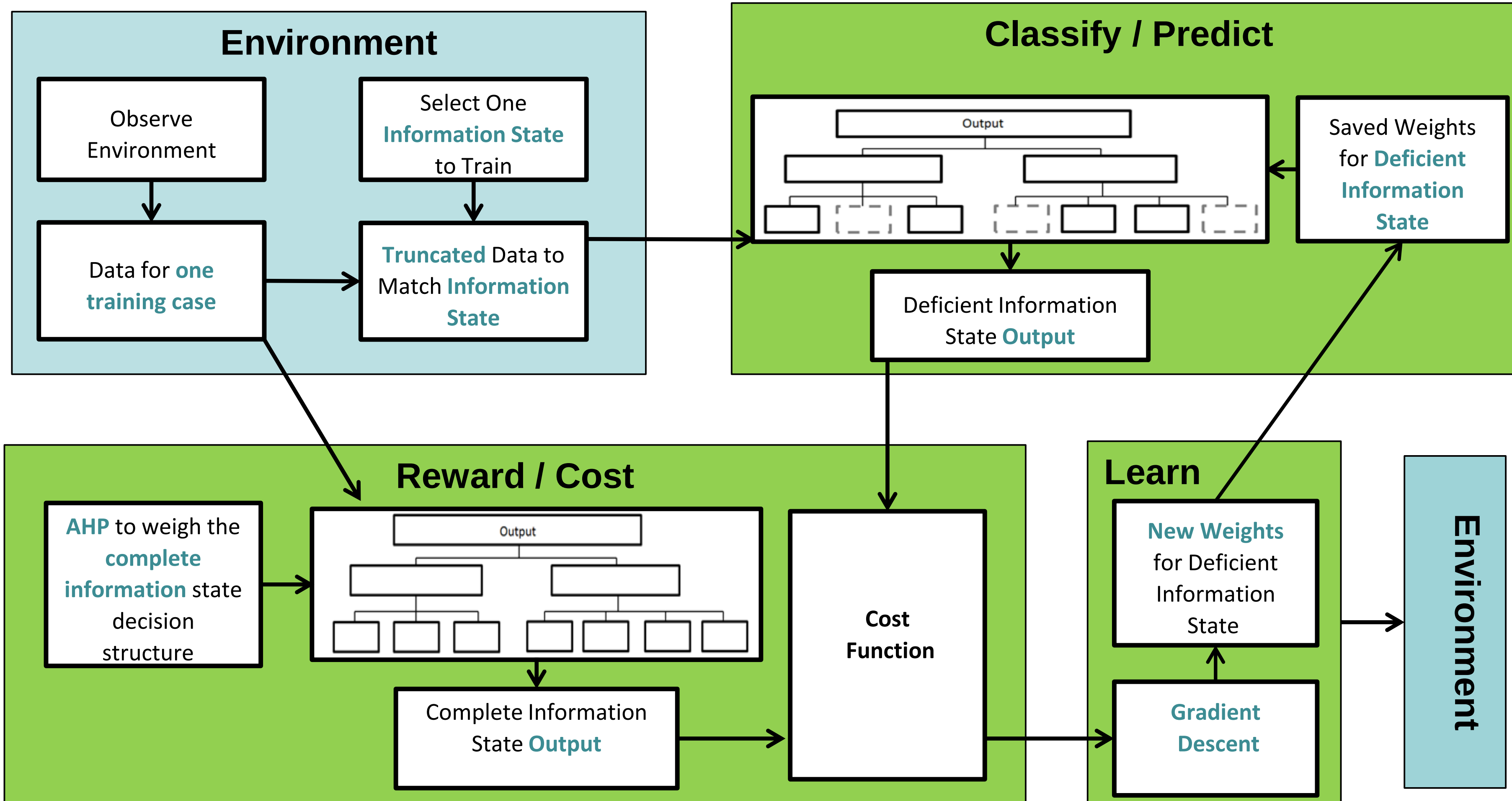
Apply the **machine learning** process to each of the **information states**

Once all of the **training cases** are exhausted, save the final weights in a **look up table** for use in practice

No matter which **combination of cues** are available, the **optimal weightings of the cues** to **predict** the complete information state is known.



Learning Process

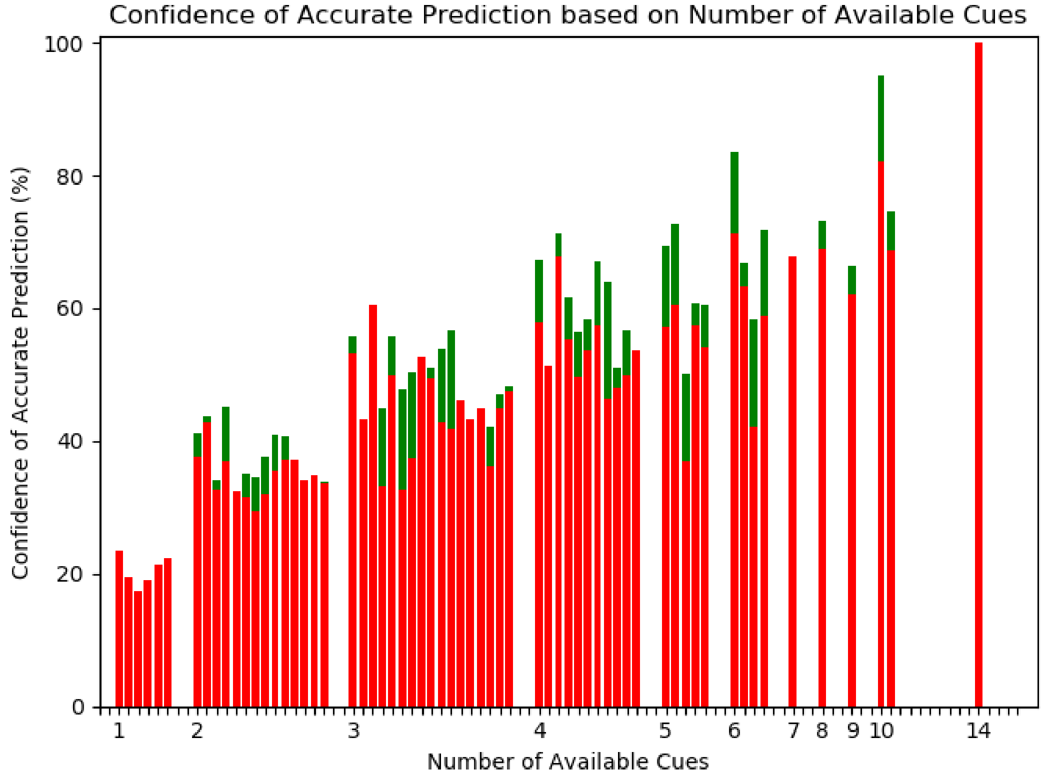


Results

A comprehensive **experiment** was performed that included **6 independent cues** and **8 dependent ones.**

Each **bar** represents and **information state**.

The **red** is the **accuracy** of the model previous to machine learning and the **green is the improvement**



	Average	Best Case	Worst Case	STD
Improvement	8 %	21%	0 %	7.9 %

Limitations

1. Inability to improve situations when the number of **available cues is low** due to **limited flexibility in weight distribution**.

2. If the **AHP generates the optimal weights from the beginning**, the learning is not required. However, it is not possible to know this without the learning process.

3. The algorithm requires access to **a several training examples** so that the weights can be optimized.

Defence Research and Development Canada
Recherche et développement pour la défense Canada

Sigmoid Supplemented Decision Structures for Evidence Sensitivity Learning

(Caelum Kamps, Rahim Jassemi-Zargani)

International Symposium on The Analytical Hierarchy Process 2018
12th – 15th July 2018

Dr. Rahim Jassemi-Zargani
Space & ISR Applications (SIA) Section
DRDC Ottawa Research Centre
Ottawa, Canada

DRDC | RDDC

Canada

Objective

The objective of this work is to address two areas in which the AHP can be supplemented to improve its reliability under specific conditions.

Those areas are:

1. **Information Deficiency**
2. **Interrelationships and Dependencies within the Cues**

The proposed solution introduces the **supplementation of sigmoid nodes** before the inputs of each weight node in the decision structures to allow for **machine learning** to adjust their **sensitivity under information deficient conditions**.

This method is beneficial because **it does not change the relative priority** derived through the **AHP** of the cues or criteria under consideration.

Sigmoid Functions

Sigmoid Functions are **monotonic** (strictly increasing or decreasing), **differentiable** (smooth), and **non-linear** functions that can be adjusted to account for relationships within the cues and output.

They have been shown to be particularly useful in **machine learning**, specifically for **neural networks**. This is partly due to their **normalization property** (bounded between 0 and 1), among other reasons.

They are composed of the following parameters: **Shape** and **Shift.**

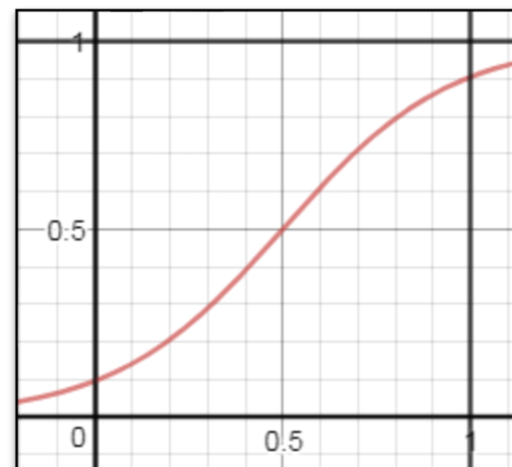
For this work we introduce the **Standard Sigmoid function (SSF)** and we initialize it to have the following **form** :

$$S(a, b, x) = \frac{1}{1+e^{-4.5(x-0.5)}}$$

Where:

- **Shift** (a) = 0.5
- **Shape** (b) = -4.5

$S(-4.5, 0.5, x)$



Sigmoid Supplemented Decision Structure

The **decision structures** for each information state are supplemented with **SSF nodes before each weight node**.

Since these SSF nodes are **very close to an identity function** on most of the interval (0,1), decisions made by the structure are **marginally impacted by the supplementation**.

$$SSF(x) \sim x \ on\ the\ interval\ (0.2,0.8)$$

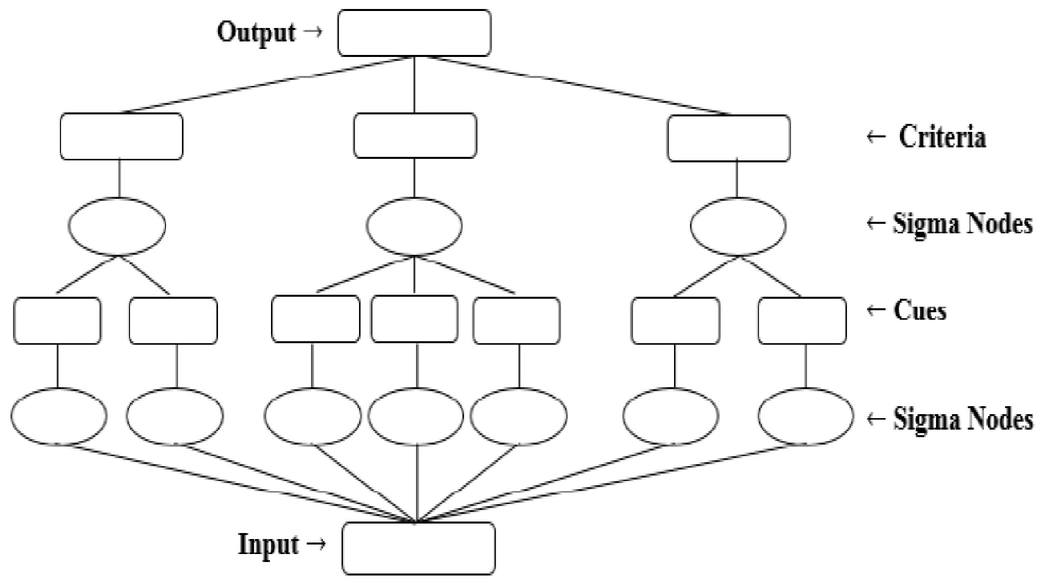


Figure 1: Sigmoid Supplemented Decision Structure

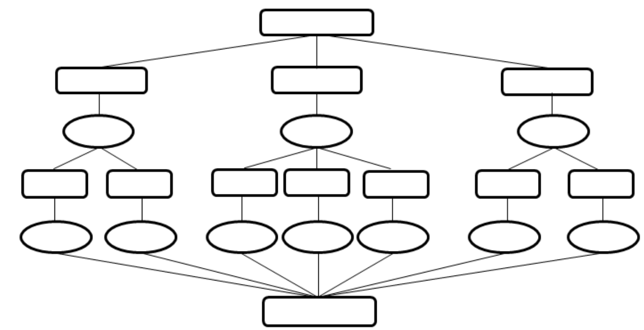
Decision Structures for Deficient Information States

There is a **decision structure** specific to each **information state.**

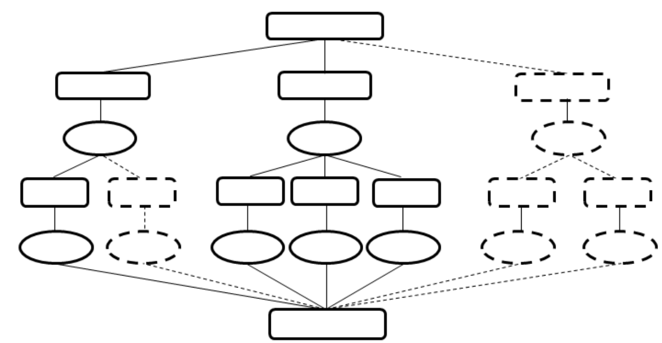
When the **information state is deficient** (missing cues) then the nodes that are impacted are **removed and the weights are recalculated** though truncated comparison matrices and the **AHP**.

Each decision structure is **initialized** with supplemented **SSF nodes.**

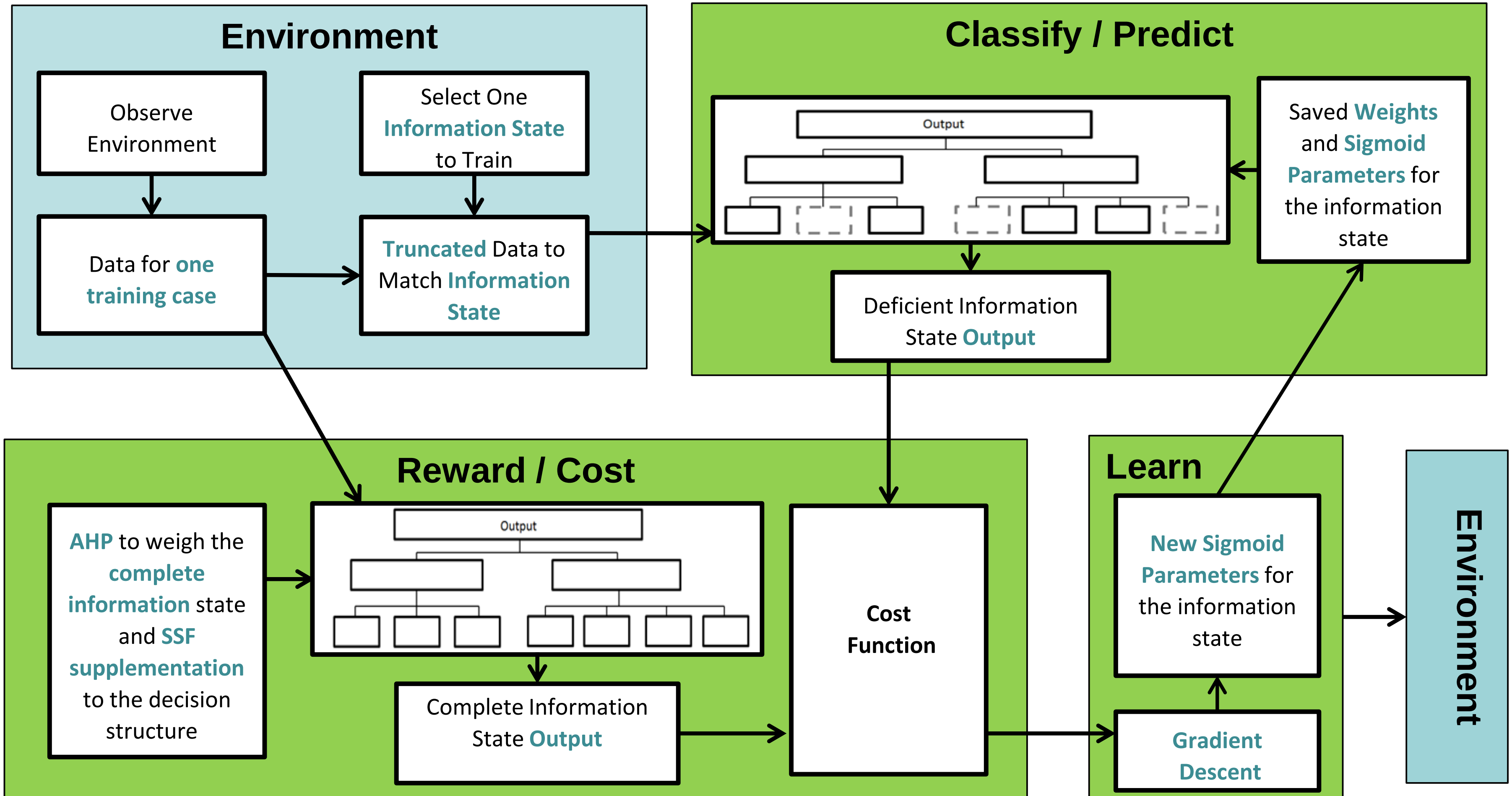
Decision Structure for the Complete Information State



Decision Structure for A Deficient Information State



Learning Process



Results

A comprehensive **experiment** was performed that included **6 independent cues** and **8 dependent ones.**

Each **bar** represents and **information state**.

The **red** is the **accuracy** of the model previous to machine learning and the **green is the improvement**

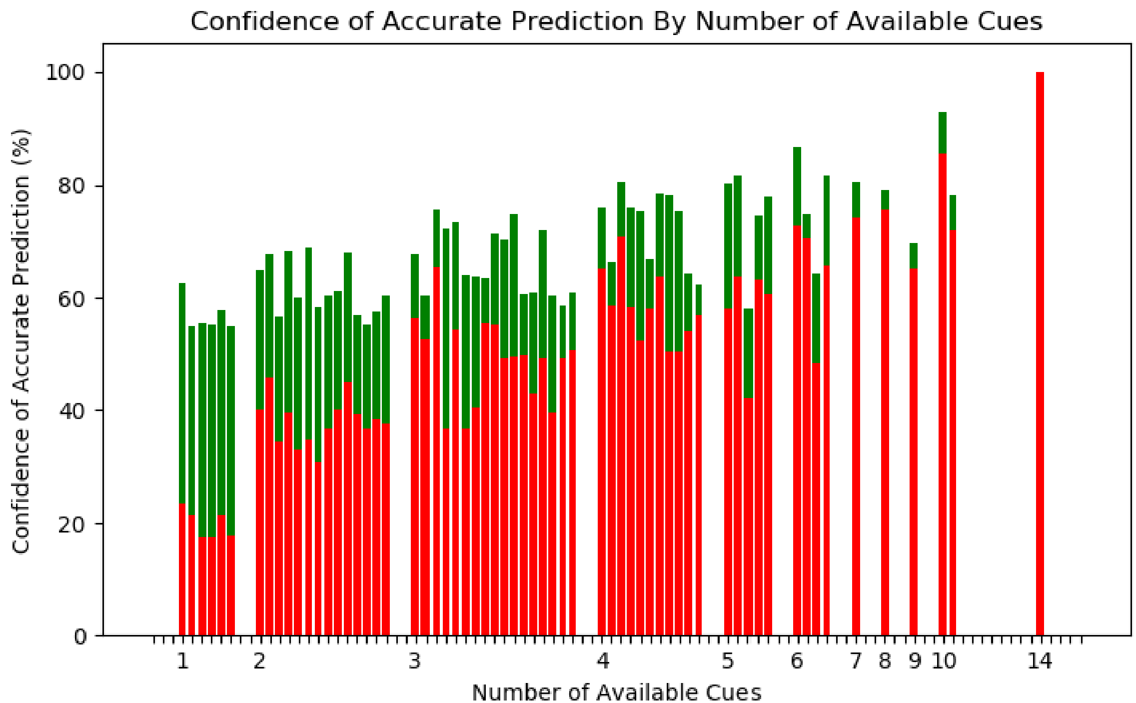


Table 1: Accurate Prediction Confidence Improvement:

	Average	Best	Worst	STD
Improvement	17.7 %	42.5 %	0.7 %	9.62 %

DRDC | RDDC

Comparison To Weight Adjustment and Limitations

Improved prediction accuracy in states of **low information availability**. This is because the **sigmoid supplemented structure** has more **flexibility**.

It is not constrained to **weight adjustment** and can even learn when only **one cue is known** using the sigma parameters.

The **sigmoid structure** can **maintain the priority ratio scales** for the **weights of the cues** created through the **AHP** process.

The method is still **limited** by the need for **training examples**.