CEDARVILLE UNIVERSITY.

Algebra of ROC Functions

Mary K. Collins[1]

[1]Cedarville University, Department of Mathematics

Introduction

Detection systems improve, save, and protect the lives of billions of people across the globe. They can be found in large scale arenas like medicine and the military but are also found in smaller and more personal areas like automobiles and home appliances. **The aim of the following research is to create a combination of detection systems that is more accurate and cost effective than any one single detection system.**

Detection Systems

The definition of a detection system is as follows: let $\mathcal{E}$ be a nonempty set representing a population of outcomes in an environment. Any $e \in \mathcal{E}$ will either be a target or non-target. Let $\mathcal{L} = \{t, n\}$ denote a label set where label t denotes a target and n denotes a non-target. It is important to note that we assume no special application for **we assume there exists a truth mapping $T : \mathcal{E} \to \mathcal{L}$ that partitions the population set $\mathcal{E}$ into $\mathcal{E}_t$ and $\mathcal{E}_n$, where $\mathcal{E}_t$ contains all target events and $\mathcal{E}_n$ contains all non-target events.** In more precise wording,

$$\mathcal{E}_t = \{e \in \mathcal{E} : T(e) = t\} \quad \text{and} \quad \mathcal{E}_n = \{e \in \mathcal{E} : T(e) = n\}.$$

The truth mapping T is unknown, so **we seek a detection system A that approximates T as close as possible.** To do this we define $\mathbf{s}$ to be a sensor that maps a stimulus from $\mathcal{E}$ to a datum in the set $\mathcal{D}$, where $\mathcal{D}$ is the set of sensor data. Since this data set may be quite large, to help filter and extract the data we use a processor $\mathbf{p}$ to refine the data. To complete the process, a classifier $\mathbf{a}_\theta$ will map a feature datum to a label in $\mathcal{L} = \{t, n\}$ for each parameter $\theta \in \Theta$, where $\Theta \subset \mathbb{R}^N$ is a set of parameters. Please note that $\mathbf{s}$, $\mathbf{p}$, and $\mathbf{a}_\theta$ are all functions. For simplicity, the entire system can be summarized graphically by,

$$\mathcal{E} \xrightarrow{\mathbf{s}} \mathcal{D} \xrightarrow{\mathbf{p}} \mathcal{F} \xrightarrow{\mathbf{a}_\theta} \mathcal{L}.$$

We define the detection system as the function composition,

$$\mathbf{A}_\theta := \mathbf{a}_\theta \circ \mathbf{p} \circ \mathbf{s} \text{ for all } \theta \in \Theta.$$

A Detection System Family (DSF) is defined by $\mathbb{A} = \{\mathbf{A}_\theta : \theta \in \Theta\}$.

Combinations

For a fixed parameter θ, a detection system combination is as follows. **We use conjunction and disjunction, more commonly known as the AND and OR rule to combine our systems.** The $\wedge$ symbol denotes the AND rule and the $\vee$ symbol denotes the OR rule. Now, for any event e in the event space $\mathcal{E}$, we have the following definitions of the above combination rules for the detection systems $\mathbb{A}$ and $\mathbb{B}$.

		$\mathbb{A}$	
	$\wedge$	t	n
$\mathbb{B}$	t	t	n
	n	n	n

		$\mathbb{A}$	
	$\vee$	t	n
$\mathbb{B}$	t	t	t
	n	t	n

Detection system family combination becomes more complicated than just detection system combination. Since fixing θ does not allow us to find the best detection system combination **we will instead use probability theory and show a formula in order to combine detection system families and ROC functions.**

ROC Functions

We require a method to assess the accuracy of classifiers; i.e. a way to quantify how closely the detection system matches the truth mapping. To this end, **we will use ROC curves, which measure the accuracy of a classifier by using the probability a classifier correctly identifies a target as a target (true positive) and the probability a classifier incorrectly identifies a non-target as a target (false positive).** We mathematically define a ROC function to be as follows. Let $\mathbb{A} = \{\mathbf{A}_\theta : \theta \in \Theta\}$ be a family of detection systems defined on the probability space $(\mathcal{E}, \mathfrak{E}, P)$ mapping to the label set $\mathcal{L} = \{t, n\}$ with parameter set Θ. For each $p \in [0, 1]$, define the set,

$$\Theta_p = \{\theta \in \Theta : P_{FP}(A_\theta) \le p\}.$$

For $p \in [0, 1]$, if Θ_p is nonempty then define,

$$f_{\mathbb{A}}(p) = \max\{P_{TP}(A_\theta) : \theta \in \Theta, P_{FP}(A_\theta) \le p\},$$

where P_{TP} is the probability of a true positive and P_{FP} is the probability of a false positive. If Θ_p is empty then $f_{\mathbb{A}}(p)$ is not defined.

Fusion Rules

Previously, we considered one detection system family, $\mathbb{A} = \{A_\theta : \theta \in \Theta\}$. Now, in order to use fusion rules, we introduce a second detection system family. Let $\mathbb{B}$ be a detection system family where we have some set of parameters Φ and $\mathbb{B} = \{B_\phi : \phi \in \Phi\}$. Then **we have the detection system families $\mathbb{A}$ and $\mathbb{B}$. We consider two main fusion rules: the AND rule and the OR rule.** The AND rule denoted with a $\wedge$ symbol and the OR denoted with a $\vee$ symbol, are defined in the following tables [3].

Let $\mathbf{C}_{(\theta,\phi)} = \mathbf{A}_\theta \wedge \mathbf{B}_\phi$. Then for any $e \in \mathcal{E}$:

	$A_\theta(e) = t$	$A_\theta(e) = n$
$B_\phi(e) = t$	$C_{\theta,\phi}(e) = t$	$C_{\theta,\phi}(e) = n$
$B_\phi(e) = n$	$C_{\theta,\phi}(e) = n$	$C_{\theta,\phi}(e) = n$

Fusion Rules (Continued)

And let $\mathbf{D}_{(\theta,\phi)} = \mathbf{A}_\theta \vee \mathbf{B}_\phi$. Then for any $e \in \mathcal{E}$:

	$A_\theta(e) = t$	$A_\theta(e) = n$
$B_\phi(e) = t$	$D_{\theta,\phi}(e) = t$	$D_{\theta,\phi}(e) = t$
$B_\phi(e) = n$	$D_{\theta,\phi}(e) = t$	$D_{\theta,\phi}(e) = n$

Now to create a formula for the AND and OR rule. For the following definitions of the AND rule and the OR rule, let $\mathbb{A}$ and $\mathbb{B}$ be two detection system families defined on $(\mathcal{E}, \mathfrak{E}, P)$ with $\mathcal{L} = \{t, n\}$. Let $\mathbf{f} = f_{\mathbb{A}}$ and $\mathbf{g} = g_{\mathbb{B}}$ be the corresponding ROC functions and let $\mathbf{f} \sqcap \mathbf{g}(r) = f_{\mathbb{A}\wedge\mathbb{B}}(r)$ and $\mathbf{f} \sqcup \mathbf{g}(r) = f_{\mathbb{A}\vee\mathbb{B}}(r)$, assume $\mathbb{A}$ and $\mathbb{B}$ are independent detection system families.

For the AND rule,

$$[\mathbf{f} \sqcap \mathbf{g}](r) = \max_{\substack{p,q \in [0,1] \\ pq = r}} \mathbf{f}(p)\mathbf{g}(q).$$

For the OR rule,

$$[\mathbf{f} \sqcup \mathbf{g}](r) = \max_{\substack{p,q \in [0,1] \\ p+q-pq = r}} \mathbf{f}(p) + \mathbf{g}(q) - \mathbf{f}(p)\mathbf{g}(q).$$

Properties

Our research has shown that **a ROC function forms a Near-Lattice under the fusion rules listed above.** We define a near-lattice as an algebraic structure that fulfills lattice criteria, with the **exception** of the **idempotent** and **absorptive** properties. The properties of a Lattice are listed below [2]

1. Closure property with respect to $\sqcap$ and $\sqcup$
2. Idempotent Property: $\mathbf{f} \sqcap \mathbf{f} = \mathbf{f}$ and $\mathbf{f} \sqcup \mathbf{f} = \mathbf{f}$
3. Absorptive Property: $\mathbf{f} \sqcap (\mathbf{f} \sqcup \mathbf{g}) = \mathbf{f}$ and $\mathbf{f} \sqcup (\mathbf{f} \sqcap \mathbf{g}) = \mathbf{f}$
4. Associative Property: $\mathbf{f} \sqcup (\mathbf{g} \sqcup \mathbf{h}) = (\mathbf{f} \sqcup \mathbf{g}) \sqcup \mathbf{h}$ and $\mathbf{f} \sqcap (\mathbf{g} \sqcap \mathbf{h}) = (\mathbf{f} \sqcap \mathbf{g}) \sqcap \mathbf{h}$
5. Commutative Property: $\mathbf{f} \sqcup \mathbf{g} = \mathbf{g} \sqcup \mathbf{f}$ and $\mathbf{f} \sqcap \mathbf{g} = \mathbf{g} \sqcap \mathbf{f}$

Example

Suppose $\mathbb{A} \in \mathrm{DSF}(\mathcal{E}, \mathcal{L})$ is mutually independent, and its ROC function is $f_{\mathbb{A}}(p) = \tanh(4p)$. The fused ROC function using the AND rule is $f_{\mathbb{A}\wedge\mathbb{A}}(p) = \left(\tanh\left(4p^{1/2}\right)\right)^2$ [1].

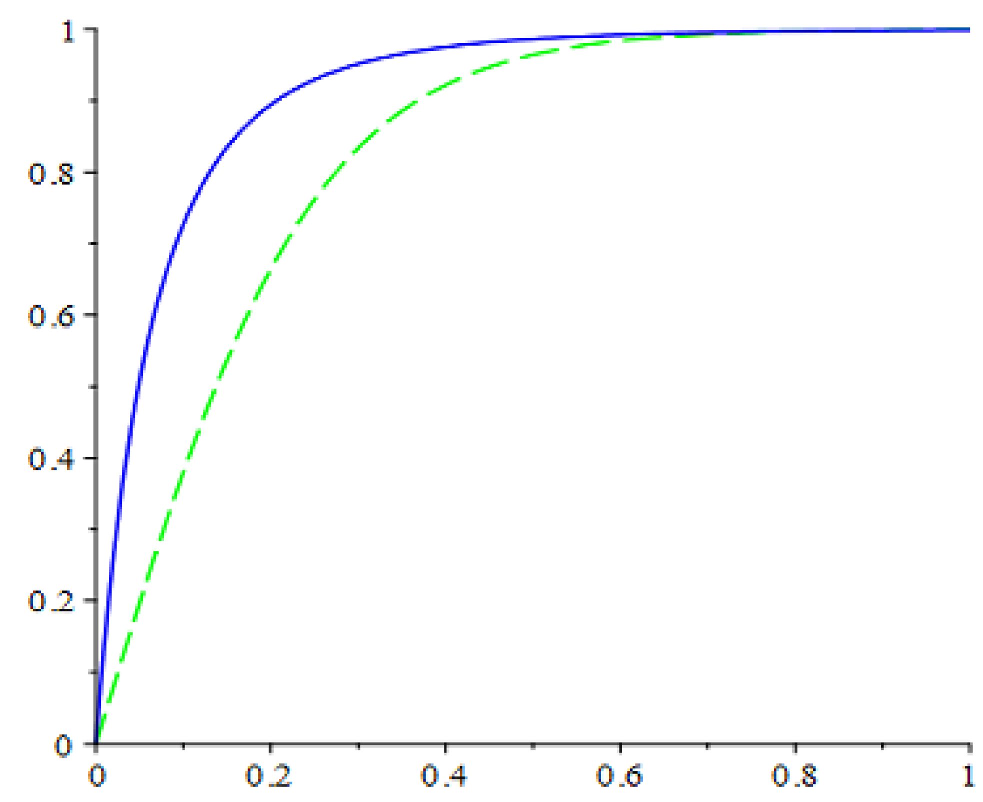

Figure 1: ROC curves of $f_{\mathbb{A}}(p) = tanh(4p)$ (dash, green), $f_{\mathbb{A}\wedge\mathbb{A}}$ (solid, blue).

Conclusion

Knowing the algebraic properties of ROC functions **will allow us to perform combinations of detection systems without individually testing each new system.** This will save organizations both time and money. Which will allow faster and better results in areas such as medicine, defense, and consumer products.

References

- Mark E. Oxley and Christine M. Schubert. "Fusion within a Detection System Family". In: *Proceedings of the 22nd International Conference on Information Fusion*. 2019.
- D.E. Rutherford. *Introduction to Lattice Theory*. New York: Hafner Publishing Company, 1965.
- C. M. Schubert. "Quantifying Correlation and its Effects on System Performance in Classifier Fusion". PhD thesis. Wright-Patterson AFB OH: AFIT/DS/ENC/05-01 Air Force Institute of Technology, 2005.

I would like to thank Dr. Mark Oxley, Dr. Christine Schubert and Ms.Schantz for their work on this research project. I would also like to thank the Air Force Institute of Technology (AFIT) for allowing me the opportunity to research and learn the different areas of mathematics.