Communications in Information and Systems
Volume 16, Number 1, 1–15, 2016

The role of prior in optimal team decisions for pattern recognition

Krishnamoorthy Kalyanam* and Meir Pachter†

Optimal team decision making subject to error-prone team members with different capabilities has been studied extensively — particularly in the context of binary classification. The over-arching goal is to correctly classify an object as either being a True or a False Target. Each team member with known Type I and II error rates is asked whether or not he determines the object to be a True Target. Based on the members' responses, a group decision is made about the identity of the object. We are interested in the optimal team decision rule that results in the least error rate or probability of misclassification. This is a widely researched topic, having applications in pattern recognition, organizational decision making, social (dichotomous) choice situations, reliability studies etc.; however, the obvious connection to information theory is missing. In this work, we establish the optimal team decision rules by direct application of Bayes decision theory. In doing so, we bring out the key role played by the parameter α that represents the known a priori probability that the object is a True Target. In particular, for a homogeneous team composition, we establish the criteria whence a majority voting scheme is optimal. Whereupon, it immediately follows that the higher the prior, α, the fewer the number of affirmative votes needed to classify the object as a True Target.

2000 Mathematics Subject Classification: Primary 00K00, 00K01; Secondary 00K02.

Key words and phrases: Pattern recognition, maximum a posteriori estimate, majority voting, dichotomous choice.

*Air Force Research Laboratory (AFRL/RQQA contractor).
†Air Force Institute of Technology (AFIT/ENG).

1. Introduction

The problem considered herein is motivated by the following operational (military) scenario. A camera equipped UAV is tasked with sequentially overflying geo-located objects of interest which need to be inspected; streaming video or photo images of the object will be transmitted to a remotely located operator or team of operators. Upon seeing the streaming video and/or photo imagery, the operator declares the inspected object of interest to be a Target (T) or, alternatively, a False Target (F). The classification decision of the operator is critical in that correct identification leads to additional assets (ground forces etc.) being assigned to engage the Target. In the same vein, allocation of assets to a False Target leads to waste of resources. So, it is imperative that the misclassification rate of the mixed initiative human-machine system be at a minimum. Towards this end, a binary classification task is considered, where an object x is inspected by a heterogenous team consisting of n members. The inspected object is either a True Target, T or a False Target, F, i.e., $x \in \{T, F\}$. In the context of military operations, a team member could represent a human operator or an Automatic Target Recognition (ATR) module. As mentioned earlier, the team could represent different operators or the same operator looking at the object from different UAV states - altitude, viewing angle and so on. For more details on the motivating operational scenario behind this work, see [2, 4]. The a priori probability that the object is a True Target, $P\{x = T\} = \alpha$, where $0 < \alpha < 1$. Henceforth, we shall refer to α as simply the prior. Member i's decision skill is parameterized by the probabilities p_i and q_i of correctly classifying True and False Targets respectively, where $0 < p_i, q_i < 1$. In general, $p_i \neq q_i$ and so, member i treats True and False Targets differently. When $p_i = q_i$, we shall say that member i is *unbiased* in that he has no bias towards either a True or a False Target. Each member i is asked whether the object is a True Target and each of their *independent* responses, $y_i \in \{Y, N\}$, $i = 1, \ldots, n$, is either in the affirmative (Y) or the negative (N). Given the set of all member responses: $\Omega = \{Y, N\}^n$, we are interested in the optimal decision rule i.e., a mapping $f : \Omega \to \{T, F\}$ that minimizes the probability of misclassification.

The problem considered herein concerns the optimal aggregation of individual judgements in dichotomous choice situations and as such has attracted considerable attention in the literature. Indeed, this problem has application in varied fields such as social choice and economic decision theory [1, 8, 10, 12], jury systems [5], electronic systems reliability [9, 11] etc. In legal parlance, the problem takes a different form wherein a jury votes on whether a defendant is innocent or guilty of committing a crime. The a priori probability that the defendant is guilty equals α. Since each member/juror is fallible with Type I and Type II errors, one is concerned with optimal decision rules that maximizes the probability that the jury will make the correct judgement. We are motivated by a fundamental result in collective decision making, the Condorcet jury theorem, which demonstrated that a group of jurists outperform a single judge [6]. In particular, the theorem shows that for the special case of $\alpha = 0.5$ and $p_i = q_i = p > 0.5,\ i = 1, \ldots, n$:

1) The probability that the committee's majority will make the right decision is higher than p and
2) The probability that the group reaches the correct decision, based on a simple majority rule, approaches 1 as $n \to \infty$.

The analysis presented here is relevant to the study of the performance, and the design, of human organizations making collective decisions. In decision theory, the problem concerns a committee with n members that accepts or rejects a project. The goal is to accept a good project and reject a bad one (with pre-specified a priori probability of a project being good equaling α). In the past, more general performance criteria have been considered wherein different costs/payoffs have been associated with the four possible outcomes i.e., accepting/ rejecting a good/ bad project [1]. In our work, we minimize the *error rate* which translates to assigning a cost of 1 for selecting a bad project and the rejection of a good project, with the other two payoffs being set to 0. In addition to being the most relevant performance measure for pattern recognition, this cost structure has the additional advantage of rendering an immediate solution by application of Bayes decision theory (see Chapter 2 in [3]). Thus, in Section 2, the minimization of the team's error rate is discussed and optimal decision rules are developed. In particular, the importance of the prior α is highlighted by considering the special case of *one* and *two* member teams in Sec. 2.2 and Sec. 2.3 respectively. Simplifications that arise when the team is homogenous and conditions whence a majority voting scheme is optimal are outlined in Section 2.4. In Section 2.5, we discuss the scenario wherein each member is *unbiased* and conditions where a dominating member's decision is optimal. Finally, some concluding remarks are presented in Section 3.

2. Error rate: probability of misclassification

We make the following standard assumption on the operator's Type I and Type II error rates:

Assumption 1.

$$p_i > 1 - q_i, \ i = 1, \ldots, n. \tag{1}$$

Remark 1. The above assumption implies that a member is more likely to correctly classify a True Target than misclassify a False Target. Also, when the prior $\alpha = 0.5$, the probability of correct classification, $p_i\alpha + q_i(1-\alpha) > 0.5$ i.e., it is better than a random guess, which is intuitively appealing.

If a team member i replies in the affirmative, we have the a posteriori probability that the object is indeed a True Target given by:

$$\bar{\alpha}_i = P\{x = T | y_i = Y\} = \frac{\alpha p_i}{\alpha p_i + (1-\alpha)(1-q_i)}. \tag{2}$$

Conversely, if a member i replies in the negative, we have the a posteriori probability that the object is a True Target given by:

$$\underline{\alpha}_i = P\{x = T | y_i = N\} = \frac{\alpha(1-p_i)}{\alpha(1-p_i) + (1-\alpha)q_i}. \tag{3}$$

From Assumption 1, it immediately follows that $\underline{\alpha}_i < \alpha < \bar{\alpha}_i$:

Lemma 1. *If Assumption 1 holds, then:*

$$\underline{\alpha}_i < \alpha < \bar{\alpha}_i. \tag{4}$$

Proof. We have:

$$\begin{aligned}
&\beta_i = p_i + q_i - 1 > 0 \\
&\Rightarrow \alpha\beta_i > \alpha^2\beta_i, \ \text{since } \alpha < 1, \\
&\Rightarrow \alpha\beta_i + \alpha(1-q_i) > \alpha^2\beta_i + \alpha(1-q_i), \\
&\Rightarrow \bar{\alpha}_i = \frac{\alpha\beta_i + \alpha(1-q_i)}{\alpha\beta_i + (1-q_i)} > \alpha. \qquad (5)
\end{aligned}$$

A similar argument shows that $\underline{\alpha}_i < \alpha$. □

Remark 2. Assumption 1 implies that the member i is reliable in that his response nudges the a posteriori probability in the right direction.

We are interested in optimal decision rules that minimize the error rate i.e., the probability of misclassification. Let the vector of classifier (team member) responses: $y = (y_1, \dots, y_n)$. Given the set of all team member responses: $\Omega = \{Y, N\}^n$, let the decision rule $f(y)$ be a mapping $f : \Omega \to \{T, F\}$. The probability of misclassification (also referred to as the error rate) associated with the rule f is given by:

$$P_E(f) = \alpha \sum_{y \in \Omega; f(y)=F} P\{y|x = T\} + (1 - \alpha) \sum_{y \in \Omega; f(y)=T} P\{y|x = F\}. \tag{6}$$

For a given $y \in \Omega$, the optimal decision, $f^*(y)$, that minimizes the error rate is given by:

$$f^*(y) = \begin{cases} T, & \text{if } P\{x = T|y\} > P\{x = F|y\}, \\ F, & \text{otherwise,} \end{cases} \tag{7}$$

where the a posteriori probabilities:

$$P\{x = T|y\} = \frac{\alpha P\{y|x = T\}}{P\{y\}} \quad \text{and} \quad P\{x = F|y\} = \frac{(1-\alpha) P\{y|x = F\}}{P\{y\}},$$

and the joint probabilities:

$$\begin{aligned} P\{y|x = T\} &= \prod_{i; y_i = Y} p_i \prod_{i; y_i = N} (1 - p_i) \\ \text{and} \quad P\{y|x = F\} &= \prod_{i; y_i = N} q_i \prod_{i; y_i = Y} (1 - q_i). \end{aligned}$$

In other words, to minimize the expected probability of error, we select the $x \in \{T, F\}$ that maximizes the a posteriori probability $P\{x|y\}$ - for proof see Sec 2.4 in [3]. Hence, the optimal decision rule is also referred to as the Maximum A Posteriori (MAP) rule. Since,

$$\begin{aligned} & P\{x = T|y\} > P\{x = F|y\} \\ & \Rightarrow \alpha P\{y|x = T\} > (1 - \alpha) P\{y|x = F\} \\ & \Rightarrow \alpha > \frac{P\{y|x = F\}}{P\{y|x = T\} + P\{y|x = F\}}, \end{aligned} \tag{8}$$

the optimal decision rule can be re-written as follows:

$$f^*(y) = \begin{cases} T, & \text{if } \alpha > \gamma(y), \\ F, & \text{otherwise,} \end{cases} \tag{9}$$

where,

$$\gamma(y) = \frac{\prod_{i;y_i=N} q_i \prod_{i;y_i=Y}(1-q_i)}{\prod_{i;y_i=Y} p_i \prod_{i;y_i=N}(1-p_i) + \prod_{i;y_i=N} q_i \prod_{i;y_i=Y}(1-q_i)}. \tag{10}$$

The above result tells us that if the prior α exceeds the threshold $\gamma(y)$, the optimal decision is to classify the object as T, else it is classified as F. Furthermore, let $\beta(y)$ denote the a posteriori probability that the object is a True Target, given the observation sequence $y \in \Omega$. For a binary classification task, since,

$$P\{x = T|y\} > P\{x = F|y\} \Rightarrow P\{x = T|y\} > 0.5,$$

the optimal decision rule takes the *intuitively* appealing form:

$$f^*(y) = \begin{cases} T, & \text{if } \beta(y) > 0.5, \\ F, & \text{otherwise.} \end{cases} \tag{11}$$

In other words, the object is declared T if the a posteriori probability that it is T given y is greater than 0.5.

For the general case, the solution strategy is the following: compute the 2^n threshold values (10): $\gamma(y)$, $\forall y \in \Omega$ and place them on the real line (between 0 and 1). Having done so, we declare $f^*(y) = T$ for all y whose threshold value lies to the left of α and $f^*(y) = F$ otherwise. As mentioned earlier, the optimal rule (9), has been derived in [1]. However, the authors therein have considered a more general cost structure and the role of the prior is hidden. We have confined our attention to minimizing the error rate, the most basic and relevant metric in pattern recognition, and by direct application of Bayes' decision theory, arrived at a simple and intuitive result. In doing so, we bring out the crucial role played by the prior, as will be seen in the sequel. We show, in the next section, the existence of a partial ordering amongst elements in Ω, which in turn brings out a useful and insightful monotonicity property in the threshold function $\gamma(y)$.

2.1. Partial ordering

Let $y \in \Omega$ such that $y_i = N$ for some $i = 1, \dots, n$. We say $z < y$ if $z \in \Omega$ such that $z \neq y$ and $z_i = Y$ for all i such that $y_i = Y$. This partial ordering gives us a monotonicity property on the threshold values (10):

Lemma 2. *$\gamma(z) < \gamma(y)$ if $z < y$.*

Proof. We will show that if any member who responded in the negative changes his vote to the affirmative (all other votes being the same), the corresponding threshold value goes down. So, let $y, z \in \Omega$ such that $z_i = y_i,\ i = 1, \ldots, j-1, j+1, \ldots, n,\ y_j = N$ and $z_j = Y$. In other words, member j changed his vote from the negative to the affirmative. So, $z < y$.

$$\begin{aligned}\gamma(y) &= \frac{\prod_{i;y_i=N} q_i \prod_{i;y_i=Y}(1-q_i)}{\prod_{i;y_i=N} q_i \prod_{i;y_i=Y}(1-q_i) + \prod_{i;y_i=Y} p_i \prod_{i;y_i=N}(1-p_i)}, \\ &= \frac{q_j A}{q_j A + (1-p_j)B},\end{aligned}$$

where, $A = \prod_{i;y_i=N;i\neq j} q_i \prod_{i;y_i=Y}(1-q_i) > 0$ and $B = \prod_{i;y_i=Y} p_i \prod_{i;y_i=N;i\neq j}(1-p_i)$. Since $p_j > 1 - q_j$, we can write:

$$\begin{aligned}&\frac{p_j}{1-q_j} > 1 > \frac{1-p_j}{q_j} \\ \Rightarrow \gamma(y) &= \frac{1}{1 + \frac{(1-p_j)B}{q_j A}} > \frac{1}{1 + \frac{p_j B}{(1-q_j)A}} \\ &= \frac{(1-q_j)A}{(1-q_j)A + p_j B} = \gamma(z).\end{aligned}$$

The last equality follows since

$$\begin{aligned}(1-q_j)A &= (1-q_j) \prod_{i;y_i=N;i\neq j} q_i \prod_{i;y_i=Y}(1-q_i) \\ &= \prod_{i;z_i=N} q_i \prod_{i;z_i=Y}(1-q_i) \\ \text{and}\quad p_j B &= p_j \prod_{i;y_i=Y} p_i \prod_{i;y_i=N;i\neq j}(1-p_i) \\ &= \prod_{i;z_i=Y} p_i \prod_{i;z_i=N}(1-p_i).\end{aligned}$$

□

The immediate implication of the monotonicity property is that if the prior satisfies: $\alpha > \gamma(y)$, then the optimal decision rule satisfies: $f^*(z) = T,\ \forall z < y$. In other words, for a given vector of member responses, y, suppose the optimal decision is to declare the object to be T. Subsequently, if any of the members who voted in the negative change their vote, the optimal decision remains unchanged. This again is an intuitively appealing result since more members voting in the affirmative makes it more likely that the object is a True Target.

To illustrate the usefulness of Lemma 2, we shall look at special cases of the model and characterize how the optimal rule changes as a function of y and its relationship with the prior. In particular, we shall examine the homogenous case, where team members are indistinguishable; thereby making a majority voting scheme relevant.

2.2. Single member: $n = 1$

For this case, there are two possible values that y can take: $y \in \{Y, N\}$. The corresponding threshold values are given by:

$$\begin{aligned}\gamma(N) &= \frac{P\{y = N|x = F\}}{P\{y = N|x = T\} + P\{y = N|x = F\}} = \frac{q_1}{q_1 + 1 - p_1},\\ \text{and} \quad \gamma(Y) &= \frac{P\{y = Y|x = F\}}{P\{y = Y|x = T\} + P\{y = Y|x = F\}} = \frac{1 - q_1}{1 - q_1 + p_1}.\end{aligned}$$

Note that Assumption 1 gives us:

$$\begin{aligned}& p_1 > 1 - q_1,\\ \Rightarrow\ & 1 - q_1 + p_1 > 2(1 - q_1),\\ \Rightarrow\ & 0.5 > \frac{1 - q_1}{1 - q_1 + p_1} = \gamma(Y).\end{aligned}$$

In a similar fashion, one can show that: $\gamma(N) > 0.5$. So, as expected from the monotonicity result (see Lemma 2), we have:

$$0 < \gamma(Y) < 0.5 < \gamma(N) < 1. \tag{12}$$

So, there is a natural progression for declaring the object to be a True Target. Indeed, if the prior is high enough i.e., $\alpha > \gamma(N)$, then we ignore the single member's response and always declare it to be T. If $\gamma(Y) < \alpha \leq \gamma(N)$, then we abide by the member's response and declare it to be T iff the member's response is in the affirmative. On the other extreme, if the prior is very low, i.e., $\alpha \leq \gamma(Y)$, then we declare the object to be F regardless of the member's response.

Remark 3. When $\alpha = 0.5$ i.e., True and False Targets are equally likely, the optimal error rate minimizing decision is to simply abide by the member's response.

2.3. Two member team: $n = 2$

For this case, there are four possible values that y can take: (N, N), (N, Y), (Y, N) and (Y, Y). We wish to compute the optimal decision rule for each outcome. As before, the monotonicity result (Lemma 2) gives us:

$$\gamma(Y, Y) < \gamma(Y, N) < \gamma(N, N) \quad \text{and} \quad \gamma(Y, Y) < \gamma(N, Y) < \gamma(N, N).$$

However, it is not clear which of $\gamma(Y, N)$ and $\gamma(N, Y)$ is greater. In other words, when the two members are in disagreement, which of the two is more likely to be correct? To address this issue, we employ the following ordering scheme.

Lemma 3. *$\gamma(Y, N) < \gamma(N, Y)$ iff $\frac{p_1 q_1}{(1-p_1)(1-q_1)} > \frac{p_2 q_2}{(1-p_2)(1-q_2)}$.*

Proof.

$$\begin{aligned}
&\gamma(Y, N) < \gamma(N, Y), \\
\Rightarrow\ & \frac{q_2(1-q_1)}{p_1(1-p_2) + q_2(1-q_1)} < \frac{q_1(1-q_2)}{p_2(1-p_1) + q_1(1-q_2)}, \\
\Rightarrow\ & q_2(1-q_1)p_2(1-p_1) < q_1(1-q_2)p_1(1-p_2), \\
\Rightarrow\ & \frac{p_2 q_2}{(1-p_2)(1-q_2)} < \frac{p_1 q_1}{(1-p_1)(1-q_1)}.
\end{aligned}$$

The proof in the other direction can be obtained by reversing the above steps. □

Without loss of generality, we order the two members such that:

$$\frac{p_1 q_1}{(1-p_1)(1-q_1)} > \frac{p_2 q_2}{(1-p_2)(1-q_2)}. \tag{13}$$

So, in lieu of Lemma 3, we can write:

$$\gamma(Y, Y) < \gamma(Y, N) < \gamma(N, Y) < \gamma(N, N). \tag{14}$$

As before, there is a natural progression for declaring the object to be a True Target. Indeed, if the prior is high enough i.e., $\alpha > \gamma(N, N)$, we ignore both members responses and declare it to be T. If $\gamma(N, Y) < \alpha \leq \gamma(N, N)$, then we declare it to be T if either member's response is in the affirmative. If $\gamma(Y, N) < \alpha \leq \gamma(N, Y)$, then we declare it to be T only if the 1^{st} (and more reliable) member's response is in the affirmative. If $\gamma(Y, Y) < \alpha \leq \gamma(Y, N)$, then we declare it to be T only if both members respond in the affirmative. Finally, at the other extreme, if $\alpha \leq \gamma(Y,Y)$, then we declare the object to be F, regardless of either member's response.

2.4. Homogenous team composition

This is the perhaps the most interesting and well studied case [7, 9–11], where $p_i = p, q_i = q, \forall i = 1,\dots,n$. Since the members are indistinguishable, it only matters as to how many of them vote in the affirmative. So, let the number of affirmative votes be denoted by $z \in \{0,\dots,n\}$. We have the joint probabilities:

$$P\{z=k|x=T\} = \begin{pmatrix} n \\ k \end{pmatrix} p^k(1-p)^{(n-k)} \text{ and}$$
$$P\{z=k|x=F\} = \begin{pmatrix} n \\ k \end{pmatrix} q^{(n-k)}(1-q)^k.$$

The optimal decision rule is given by:

$$f^*(k) = \begin{cases} T, & \text{if } \alpha > \gamma(k), \\ F, & \text{otherwise}, \end{cases} \tag{15}$$

where, as before, the threshold corresponding to k members voting in the affirmative is given by:

$$\begin{aligned} \gamma(k) &= \frac{P\{z=k|x=F\}}{P\{z=k|x=T\}+P\{z=k|x=F\}} \\ &= \frac{q^{(n-k)}(1-q)^k}{q^{(n-k)}(1-q)^k + p^k(1-p)^{(n-k)}}. \end{aligned} \tag{16}$$

For the homogenous team composition, we can do better than the partial ordering result available for the general case.

Lemma 4. *From Assumption 1, we get the full ordering:*

$$\gamma(n) < \cdots < \gamma(0). \tag{17}$$

Proof. From Assumption 1, we have:

$$\begin{aligned} &1-p-q<0, \\ \Rightarrow\ &(1-p)(1-q) < pq. \end{aligned}$$

Multiplying both sides by $((1-p)q)^{n-k-1}(p(1-q))^k$, we get:

$$(1-p)^{n-k}(1-q)^{k+1}q^{n-k-1}p^k < p^{k+1}q^{n-k}(1-p)^{n-k-1}(1-q)^k.$$

Adding $q^{2(n-k)-1}(1-q)^{2k+1}$to both sides, we get:

$$\begin{aligned}
&q^{n-k-1}(1-q)^{k+1}\left[q^{n-k}(1-q)^k+(1-p)^{n-k}p^k\right] \\
&< q^{n-k}(1-q)^k\left[q^{n-k-1}(1-q)^{k+1}+(1-p)^{n-k-1}p^{k+1}\right], \\
\Rightarrow\ & \frac{q^{n-k-1}(1-q)^{k+1}}{q^{n-k-1}(1-q)^{k+1}+(1-p)^{n-k-1}p^{k+1}} < \frac{q^{n-k}(1-q)^k}{q^{n-k}(1-q)^k+(1-p)^{n-k}p^k}, \\
\Rightarrow\ & \gamma(k+1) < \gamma(k).
\end{aligned}$$

□

From Lemma 4, we have: $0 < \gamma(n) < \cdots < \gamma(0) < 1$. So, the optimal *minimum* number of affirmative votes needed to declare the object to be a True Target is given by:

$$k^*(\alpha) = \begin{cases} 0, & \text{if } \alpha > \gamma(0), \\ 1, & \text{if } \gamma(0) \geq \alpha > \gamma(1), \\ \vdots & \\ n, & \text{if } \gamma(n-1) \geq \alpha > \gamma(n), \end{cases} \tag{18}$$

and, finally if $\alpha \leq \gamma(n)$, the object is declared a False Target. Suppose we have an odd number of members, i.e., $n = 2m+1$. A simple majority voting scheme is given by:

$$f_M(k) = \begin{cases} T, & \text{if } k > m, \\ F, & \text{otherwise}, \end{cases} \tag{19}$$

i.e., if at least $(m+1)$ members vote in the affirmative, then the object is declared a T. For details on application of majority voting to pattern recognition and analysis of its performance, see [7].

Corollary 1. *The simple majority voting scheme is optimal i.e.,* $k^*(\alpha) = m+1$ *iff* $\gamma(m) \geq \alpha > \gamma(m+1)$.

If in addition, the team members are unbiased, homogeneous and the prior $\alpha = 0.5$, the simple majority voting scheme is optimal as shown below.

Lemma 5. *If* $p_i = q_i = p, \forall i$ *and* $\alpha = 0.5$, *then* $\gamma(m) \geq \alpha > \gamma(m+1)$.

Proof. The threshold corresponding to k members voting in the affirmative is given by:

$$\gamma(k) = \frac{p^{(n-k)}(1-p)^k}{p^{(n-k)}(1-p)^k + p^k(1-p)^{(n-k)}} = \frac{1}{1+\left(\frac{p}{1-p}\right)^{2k-n}}.$$

Assumption 1 gives us: $2p > 1$ and so, we have:

$$\gamma(m+1) = \frac{1}{1+\left(\frac{p}{1-p}\right)} = 1-p < 0.5 \text{ and } \gamma(m) = \frac{1}{1+\left(\frac{1-p}{p}\right)} = p > 0.5.$$

Hence, $\gamma(m+1) < \alpha < \gamma(m)$. □

The above result confirms our common sense notion that when the team members are identical and have no bias, two opposing members' decisions cancel each other out and so, a simple majority rule (or *democracy*) is indeed optimal. As noted earlier, Condorcet's jury theorem [6] shows that the probability that a homogenous team comes to the correct decision (based on simple majority rule) approaches 1 as $n \to \infty$.

2.5. Unbiased team members

For this case (considered in [8]), $p_i = q_i, \forall i = 1, \ldots, n$. So, the joint probabilities:

$$P\{y|x=T\} = \prod_{i;y_i=Y} p_i \prod_{i;y_i=N} (1-p_i)$$
$$\text{and} \quad P\{y|x=F\} = \prod_{i;y_i=N} p_i \prod_{i;y_i=Y} (1-p_i).$$

The threshold value corresponding to $y \in \Omega$ is given by:

$$\gamma(y) = \frac{\prod_{i;y_i=N} p_i \prod_{i;y_i=Y}(1-p_i)}{\prod_{i;y_i=N} p_i \prod_{i;y_i=Y}(1-p_i) + \prod_{i;y_i=Y} p_i \prod_{i;y_i=N}(1-p_i)}. \tag{20}$$

Let $\bar{y}$ be the complement of y such that $\bar{y}_i = Y$ if $y_i = N$ and $\bar{y}_i = N$ if $y_i = Y$. It immediately follows that:

$$\gamma(\bar{y}) = 1 - \gamma(y). \tag{21}$$

So, the 2^n threshold values exhibit symmetry about 0.5 in the real line between 0 and 1. So, for this case, the threshold values exhibit the complementary symmetry property (21), in addition to the monotonicity property in Lemma 2. Furthermore, one can always order the members such that: $p_1 > p_2 > \cdots > p_n > 0.5$, where $p_n > 0.5$ follows from Assumption 1.

Lemma 6. *If* $\alpha = 0.5$ *and* $p_1 > \gamma(y_2, \ldots, y_n),\ y_k = N, k = 2, \ldots, n$, *the optimal decision rule:* $f^*(y) = T$ *iff* $y_1 = Y$.

Proof. We have:

$$\begin{aligned} &p_1 > \frac{\prod_{i=2}^n p_i}{\prod_{i=2}^n p_i + \prod_{i=2}^n (1-p_i)}, \\ &\Rightarrow p_1 \prod_{i=2}^n (1-p_i) > (1-p_1) \prod_{i=2}^n p_i, \\ &\Rightarrow \frac{p_1 \prod_{i=2}^n p_i}{p_1 \prod_{i=2}^n p_i + (1-p_1)\prod_{i=2}^n (1-p_i)} > 0.5 = \alpha, \\ &\Rightarrow \gamma(N, y_2, \ldots, y_n) > \alpha,\ y_k = Y, k = 2, \ldots, n. \end{aligned}$$

From the monotonicity property (Lemma 2), we have:

$$\gamma(N, y_2, \ldots, y_n) > \gamma(N, Y, \ldots, Y) > \alpha, \quad y_k \in \{Y, N\},\ k = 2, \ldots, n.$$

From the complementary symmetry property (21), we have:

$$\gamma(N, y_2, \ldots, y_n) > \alpha > \gamma(Y, y_2, \ldots, y_n), \quad y_k \in \{Y, N\},\ k = 2, \ldots, n.$$

$$\Rightarrow f^*(y) = \begin{cases} T, & \text{if } y_1 = Y, \\ F, & \text{otherwise.} \end{cases} \tag{22}$$

Hence, it is optimal to abide by member 1's response. □

In other words, if member 1 dominates the rest of the team members put together, his response is optimal. Hence, for this scenario, *autocracy* is optimal as opposed to *democracy*, which was optimal under a homogenous team setting (see Lemma 5). For instance, when $n = 2$, the condition in Lemma 6 collapses to: $p_1 > p_2$ which, by definition, is true. So, we have:

$$\gamma(Y, Y) < \gamma(Y, N) < 0.5 < \gamma(N, Y) < \gamma(N, N). \tag{23}$$

So, when the prior $\alpha = 0.5$, the optimal decision is to always agree with member 1's response. Furthermore,

$$\gamma(Y,N) = \frac{p_2(1-p_1)}{p_1+p_2-2p_1p_2} \quad \text{and} \quad \gamma(N,Y) = \frac{p_1(1-p_2)}{p_1+p_2-2p_1p_2}. \tag{24}$$

So, in the limit as $p_1 \to 1$, $\gamma(Y,N) \to 0$ and $\gamma(N,Y) \to 1$, we have $\gamma(Y,N) < \alpha < \gamma(N,Y)$ for any α. Hence, the optimal decision is to agree with member 1 for any α, when he is close to being perfect, as dictated by common sense.

3. Conclusion

We have prescribed the optimal decision rule for a team of n fallible members entrusted with the task of binary classification, by direct application of Bayes decision theory. For the most general case, when a team member is susceptible to both Type I and II errors, the problem reduces to the computation of 2^n threshold values. If a threshold value is less than the prior, then the corresponding optimal decision is to declare the object to be a True Target. Otherwise, it is declared to be a False Target. We have also established a monotonicity property on the threshold values, which comes about due to a partial ordering of the team members' responses. For the special case of a homogenous team, we recover a full ordering, thereby establishing criteria under which a majority voting rule is optimal. Conversely, for a diverse team with *unbiased* members and a prior of 0.5, we show that abiding by a dominant member's verdict is, in fact, optimal.

References

[1] R. C. Ben-Yashar and S. I. Nitzan, *The optimal decision rule for fixed-size committees in dichotomous choice situations: The general result*, International Economic Review **38** (1997), no. 175-186.

[2] P. Chandler, M. Patzek, M. Pachter, C. Rothwell, S. Naderer, and K. Kalyanam, *Integrated human behavior modeling and stochastic control (IHBMSC)*, Final Report AFRL-RQ-WP-TR-2014-0191, Air Force Research Lab (2014).

[3] R. O. Duda and P. E. Hart, *Pattern classification and scene analysis*, Wiley-Interscience (1973).

[4] K. Kalyanam, M. Pachter, M. Patzek, C. Rothwell, and S. Darbha, *Optimal human-machine teaming for a sequential inspection operation*,

IEEE Transactions on Human-Machine Systems **46** (2016), no. 4, 557–568.

[5] R. Kirstein, *The condorcet jury-theorem with two independent error-probabilities*, CSLE Discussion Paper 2006-03 (2006).

[6] K. K. Ladha, *The condorcet jury theorem, free speech, and correlated votes*, American Journal of Political Science **36** (1992), no. 3, 617–634.

[7] L. Lam and C. Y. Suen, *Application of majority voting to pattern recognition: An analysis of its behavior and performance*, IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems and Humans **27** (1997), no. 5, 553–568.

[8] S. I. Nitzan and J. Paroush, *Optimal decision rules in uncertain dichotomous choice situations*, International Economic Review **23** (1982), no. 2, 289–297.

[9] R. K. Sah, *An explicit closed-form formula for profit-maximizing k-out-of-n systems subject to two kinds of failures*, Microelectronics and Reliability **30** (1990), no. 6, 1123–1130.

[10] R. K. Sah and J. E. Stiglitz, *Committees, hierarchies and polyarchies*, The Economic Journal **98** (1988), 451–470.

[11] R. K. Sah and J. E. Stiglitz, *Qualitative properties of profit-making k-out-of-n systems subject to two kinds of failures*, IEEE Transactions on Reliability **37** (1988), no. 5, 515–521.

[12] P. Stone, *Introducing difference into the Condorcet jury theorem*, Decision Theory **78** (2015), 399–409.

InfoSciTex Corporation
4027 Colonel Glenn Hwy. Ste. 210
Dayton, OH 45431, USA
E-mail address: krishnak@ucla.edu

Department of Electrical & Computer Engineering
Air Force Institute of Technology
Wright-Patterson A.F.B., OH 45433, USA
E-mail address: meir.pachter@afit.edu