Supplementary Material for IJCAI19 Submission: Leveraging Uncertainty in Deep Learning for Selective Classification

1 Model Derivation for MIPCSC

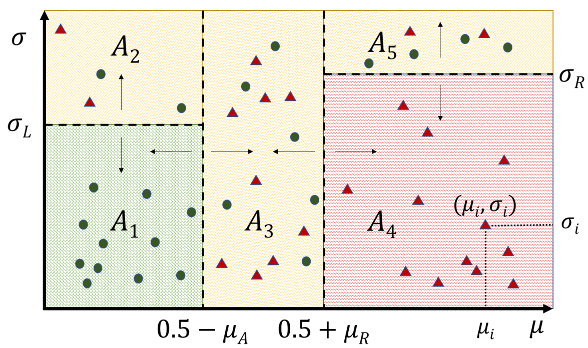

Figure 1: Graphical Illustration of the MIPSC model

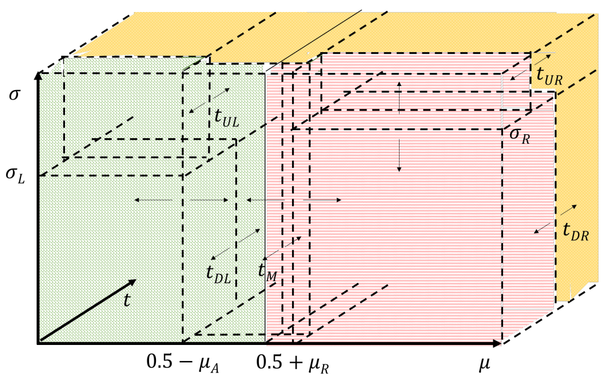

Figure 2: Graphical Illustration of the MIPCSC model

First, we inherit the following constraints from MIPSC (refer to Figure 1) to define instances to be on the regions, A_1, A_2, A_3, A_4 or A_5 before we start extend decisions to the third dimension.

The following constraint regulates the samples which do not reside in A_3 based on their predictive means but on the right hand side of A_3 such that $i \in A_4 \cup A_5$:

Variable	Definition
y_i	Ground truth label of instance i
p_i	Positive classification indicator for instance i
n_i	Negative classification indicator for instance i
r_i	Rejection indicator for instance i
μ_i	Predictive mean for instance i
σ_i	Uncertainty for instance i
μ_L	Left boundary for rejection
μ_R	Right boundary for rejection
σ_L	Upper uncertainty boundary for positive decisions
σ_R	Upper uncertainty boundary for negative decisions
L_i	Left area indicator for instance i
R_i	Right area indicator for instance i
D_{L_i}	Down-left area indicator for instance i
D_{R_i}	Down-right area indicator for instance i
$rCap$	Rejection capacity

Table 1: Notation Table for MIPSC

$$\mu_i > 0.5 + \mu_R \text{ iff } R_i = 1 \tag{1}$$

Now, we would like to distinguish the instances between A_4 and A_5. The following constraints characterize the samples that conform to A_4 such that $i \in A_4$:

$$\sigma_i < \sigma_R \text{ iff } D_{R_i} = 1 \tag{2}$$

Similarly, the following constraint define the samples which do not reside in A_3 based on their predictive means but on the left hand side of A_3 such that $i \in A_1 \cup A_2$:

$$\mu_i > 0.5 - \mu_A \text{ iff } L_i = 1 \tag{3}$$

Further, we would like to distinguish the instances between A_1 and A_2. The following constraints characterize the samples that conform to A_1 such that $i \in A_1$:

$$\sigma_i < \sigma_L \text{ iff } D_{L_i} = 1 \tag{4}$$

Now, we start extending our model to the third dimension (refer to Figure 2). The following constraint focuses on the

Variable	Definition
t_i	Value for instance i
p_{ij}	Positive classification indicator for instance i and area j
n_{ij}	Negative classification indicator for instance i and area j
r_i	Rejection indicator for instance i
t_{DL}	Down-left area value boundary for rejection
t_{UL}	Upper-left area value boundary for rejection
t_M	Middle area value boundary for rejection
t_{DR}	Down-right area value boundary for rejection
t_{UR}	Upper-right area value boundary for rejection
S_{DL_i}	Surface-down-left area indicator for instance i
S_{DR_i}	Surface-down-right area indicator for instance i
S_{UL_i}	Surface-up-left area indicator for instance i
S_{UR_i}	Surface-up-right area indicator for instance i
S_{M_i}	Surface-down-middle area indicator for instance i

Table 2: Additional Notation Table for MIPCSC

region A_1 and finds the value threshold for that region. If the transaction corresponds to A_1 region and its value is less than the region's value threshold, then our model makes a positive decision.

$$t_i < t_{DL} \text{ iff } S_{DL_i} = 1 \tag{5}$$
$$L_i + D_{L_i} + S_{DL_i} > 2 \text{ iff } p_{i1} = 1 \tag{6}$$

Now, we would like to find our decision threshold for A_2. Similarly to the previous constraints, if the transaction corresponds to A_2 region and its value is less than the region's value threshold, then our model makes a positive decision.

$$t_i < t_{UL} \text{ iff } S_{UL_i} = 1 \tag{7}$$
$$L_i + (1 - D_{L_i}) + S_{UL_i} > 2 \text{ iff } p_{i2} = 1 \tag{8}$$

Similar to the positive decision regions, now, we focus on the negative decision regions: A_4 and A_5. The following constraint focuses on the region A_4 and finds the value threshold for that region. If the transaction corresponds to A_4 region and its value is less than the region's value threshold, then our model makes a negative decision.

$$t_i < t_{DR} \text{ iff } S_{DR_i} = 1 \tag{9}$$
$$R_i + D_{R_i} + S_{DR_i} > 2 \text{ iff } n_{i1} = 1 \tag{10}$$

Now, we would like to find our decision threshold for A_5. Similarly to the previous constraints, if the transaction corresponds to A_5 region and its value is less than the region's value threshold, then our model makes a negative decision.

$$t_i < t_{UR} \text{ iff } S_{UR_i} = 1 \tag{11}$$
$$R_i + (1 - D_{R_i}) + S_{UR_i} > 2 \text{ iff } n_{i2} = 1 \tag{12}$$

Finally, we move onto our middle region, A_3. Here, we would like our model to make a positive or negative decision using the predictive mean of 0.5 as the threshold and considering the value threshold we optimally determine by solving the problem, t_M.

$$\mu_i > 0.5 \text{ iff } Q_i = 1 \tag{13}$$
$$t_i < t_M \text{ iff } S_{M_i} = 1 \tag{14}$$
$$(2 - L_i - R_i) + S_{M_i} + (1 - Q_i) > 3 \text{ iff } p_{i3} = 1 \tag{15}$$
$$(2 - L_i - R_i) + S_{M_i} + Q_i > 3 \text{ iff } n_{i3} = 1 \tag{16}$$

As we have constrained our positive and negative classification decision regions, we reject the remaining instances covered by the constraint below:

$$\sum_{j=1}^{n} [p_{ij}] + \sum_{j=1}^{n} [n_{ij}] + r_i = 1, \forall i \tag{17}$$

where the reject decision is assigned when our model cannot a make positive or negative classification decision for instance i due to DNN uncertainty, predictive mean, or it does not make financial sense to spend money on a reject decision.

Following our definition and our constraints, we propose our cost-sensitive framework called Mixed-Integer Programming based Cost-Sensitive Selective Classification (MIPCSC) formally as follows:

$$\begin{aligned} \underset{\substack{\mu_L,\mu_R,\sigma_L,\sigma_R,\\ t_{DL},t_{UL},t_M,t_{DR},t_{UR}}}{\text{maximize}} \quad & \omega_{tp}(\sum_{i=1}^{n}\sum_{j=1}^{3} p_{ij}(1-y_i)t_i + \sum_{i=1}^{n} r_i(1-y_i)t_i) \\ & +\omega_{tn}(\sum_{i=1}^{n}\sum_{j=1}^{3} n_{ij}y_it_i + \sum_{i=1}^{n} r_iy_it_i) \\ & -\omega_{fn}(\sum_{i=1}^{n}\sum_{j=1}^{3} n_{ij}(1-y_i)t_i) \\ & -\omega_{fp}(\sum_{i=1}^{n}\sum_{j=1}^{3} p_{ij}y_it_i) - c\sum_{i=1}^{m} r_i \end{aligned} \tag{18}$$

$$\mu_R - \epsilon + MR_i \geq \mu_i - 0.5 \geq \mu_R - M(1 - R_i), \forall i \tag{19}$$
$$M(1 - L_i) - \mu_L \geq \mu_i - 0.5 \geq \epsilon - \mu_L - ML_i, \forall i \tag{20}$$
$$\sigma_L + M(1 - D_{L_i}) \geq \sigma_i \geq \sigma_L + \epsilon - MD_{L_i}, \forall i \tag{21}$$
$$\sigma_R + M(1 - D_{R_i}) \geq \sigma_i \geq \sigma_R + \epsilon - MD_{R_i}, \forall i \tag{22}$$
$$0.5 + \epsilon + MQ_i \geq \mu_i \geq 0.5 + M(Q_i - 1), \forall i \tag{23}$$
$$t_{DL} + M(1 - S_{DL_i}) \geq t_i \geq t_{DL} + \epsilon - S_{DL_i}, \forall i \tag{24}$$
$$t_{UL} + M(1 - S_{UL_i}) \geq t_i \geq t_{UL} + \epsilon - S_{UL_i}, \forall i \tag{25}$$
$$t_M + M(1 - S_{M_i}) \geq t_i \geq t_M + \epsilon - S_{M_i}, \forall i \tag{26}$$
$$t_{DR} + M(1 - S_{DR_i}) \geq t_i \geq t_{DR} + \epsilon - S_{DR_i}, \forall i \tag{27}$$
$$t_{UR} + M(1 - S_{UR_i}) \geq t_i \geq t_{UR} + \epsilon - S_{UR_i}, \forall i \tag{28}$$

$$D_{L_i} + L_i + S_{DL_i} \geq 3p_{i1}, \forall i \quad (29)$$
$$D_{L_i} + L_i + S_{DL_i} - 2 \leq 3p_{i1}, \forall i \quad (30)$$
$$(1 - D_{L_i}) + L_i + S_{DL_i} \geq 3p_{i2}, \forall i \quad (31)$$
$$(1 - D_{L_i}) + L_i + S_{DL_i} - 2 \leq 3p_{i2}, \forall i \quad (32)$$
$$D_{R_i} + R_i + S_{DR_i} \geq 3n_{i1}, \forall i \quad (33)$$
$$D_{R_i} + R_i + S_{DR_i} - 2 \leq 3n_{i1}, \forall i \quad (34)$$
$$(1 - D_{R_i}) + R_i + S_{DR_i} \geq 3n_{i2}, \forall i \quad (35)$$
$$(1 - D_{R_i}) + R_i + S_{DR_i} - 2 \leq 3n_{i2}, \forall i \quad (36)$$
$$(2 - L_i - R_i) + S_{M_i} + (1 - Q_i) \geq 4p_{i3}, \forall i \quad (37)$$
$$(2 - L_i - R_i) + S_{M_i} + (1 - Q_i) - 3 \leq 4p_{i3}, \forall i \quad (38)$$
$$(2 - L_i - R_i) + S_{M_i} + Q_i \geq 4n_{i3}, \forall i \quad (39)$$
$$(2 - L_i - R_i) + S_{M_i} + Q_i - 3 \leq 4n_{i3}, \forall i \quad (40)$$
$$\sum_{j=1}^{3} [p_{ij}] + \sum_{j=1}^{3} [n_{ij}] + r_i = 1, \forall i \quad (41)$$
$$\forall p_{ij}, r_i, n_{ij}, R_i, L_i, D_{L_i}, D_{R_i}$$
$$, S_{DL_i}, S_{UL_i}, S_{M_i}, S_{DR_i}, S_{UR_i} \in \{0, 1\} \quad (42)$$
$$\forall i \in \{1...m\}, and \;\; \mu_L, \mu_R, \sigma_L, \sigma_R \in \mathbb{R} \quad (43)$$

In this formulation, constraint (19) is derived from (1), (20) is derived from (3), (21) is derived from (4), (22) is derived from (2).

Then, constraint (23) is derived from (13), (24) is derived from (5), (25) is derived from (7), (26) is derived from (14), (27) is derived from (9), and (28) is derived from (11).

Next, (29) and (30) are derived from (6), (31) and (32) are derived from (8), (33) and (34) are derived from (10), and (35) and (36) are derived from (12).

Finally, (37) and (38) are derived from (15), and (39) and (40) are derived from (16) following the Big-M method.

2 More Experimental Results

Please refer to Figure 3 for the experimental results that were reported in the main paper and refer Figure 4 for the additional ones below.

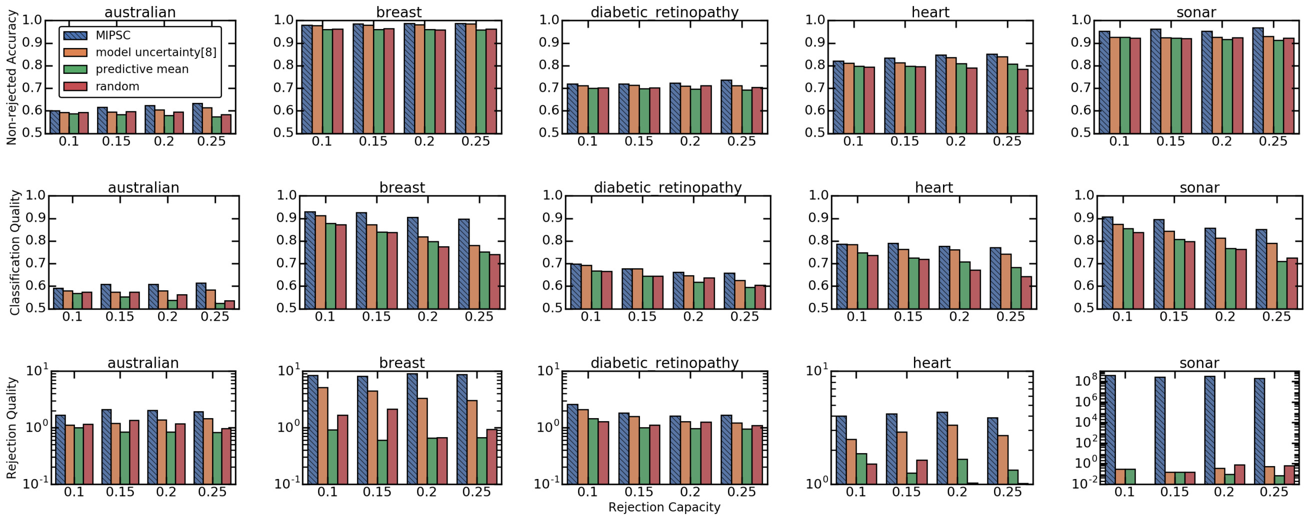

Figure 3: **Performance of the MIPSC and other baselines under varying rejection capacities. Notice the superior performance of MIPSC over the recent state-of-the-art and other baselines in all of the three performance metrics for publicly available datasets: australian, breast, diabetic, heart, sonar.**

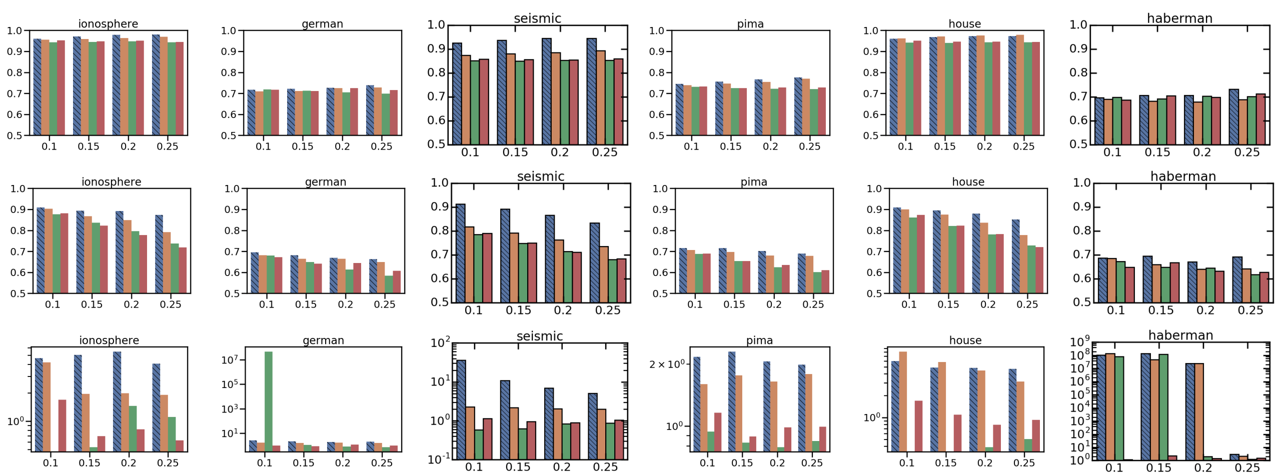

Figure 4: **Performance of the MIPSC and other baselines under varying rejection capacities. Notice the superior performance of MIPSC over the recent state-of-the-art and other baselines in all of the three performance metrics for publicly available datasets: ionosphere, german, seismic, pima, house, haberman.**