TRAINING INDICES;

Predictors	Deviance	Deviance (SE)	AIC	AIC (SE)	LRT	LRT (SE)	Wald Test	Wald Test (SE)
+ EDAD	92653.98	360.41	92657.98	360.41	0.0000	0.0000	0.0000	0.0000
+ DIABETES	91434.31	352.82	91440.31	352.82	0.0000	0.0000	0.0000	0.0000
+ SEXO	90407.51	358.89	90415.51	358.89	0.0000	0.0000	0.0000	0.0000
+ RENAL_CRONICA	89918.23	364.27	89928.23	364.27	0.0000	0.0000	0.0000	0.0000
+ OBESIDAD	89563.71	358.45	89575.71	358.45	0.0000	0.0000	0.0000	0.0000
+ HIPERTENSION	89432.67	352.01	89446.67	352.01	0.0000	0.0000	0.0000	0.0000
+ INMUSUPR	89378.93	341.04	89394.93	341.04	0.0000	0.0000	0.0000	0.0000
+ EPOC	89360.49	336.19	89379.69	336.88	0.0015	0.0011	0.0006	0.0004
+ ASMA	89362.62	344.73	89381.42	344.02	0.0017	0.0019	0.0018	0.0020
+ CARDIOVASCULAR	89353.34	339.73	89375.74	339.79	0.2160	0.2044	0.2233	0.2071
+ TABAQUISMO	89352.87	339.54	89376.47	339.49	0.2588	0.2118	0.2604	0.2103

TRAIN & TEST ERRORS;

```
require(ggplot2);
require(gridExtra);

test_error <- readRDS("~/TABASCO-MEXCOV-19/src/models/logit-deceased/freq/5-fold_Cross_Validation/test_
train_error <- readRDS("~/TABASCO-MEXCOV-19/src/models/logit-deceased/freq/5-fold_Cross_Validation/trai

# Complexity Vs (Cutoff, Accuracy, Sensitivity, Specificity) on TRAIN and TEST sets;
{
cut.plot <- ggplot() +

  # 5-fold CV Cutoff;
  geom_line(aes(x = c(1:11), y = apply(X = test_error$CUTOFF, MARGIN = 1, FUN = mean)),
            col = "#009dd0",
            size = 0.75) +
  geom_line(aes(x = c(1:11), y = apply(X = train_error$CUTOFF, MARGIN = 1, FUN = mean)),
            col = "#f58f3b",
            size = 0.75) +

  geom_ribbon(aes(x = c(1:11),
                  ymin = apply(X = test_error$CUTOFF, MARGIN = 1, FUN = min),
                  ymax = apply(X = test_error$CUTOFF, MARGIN = 1, FUN = max)),
              alpha    = 0.1,
              linetype = "dashed",
              colour   = "#009dd0",
              size     = 0.75,
              fill     = "#009dd0") +
  geom_ribbon(aes(x = c(1:11),
                  ymin = apply(X = train_error$CUTOFF, MARGIN = 1, FUN = min),
                  ymax = apply(X = train_error$CUTOFF, MARGIN = 1, FUN = max)),
              alpha    = 0.1,
              linetype = "dashed",
              colour   = "#f58f3b",
              size     = 0.75,
```

```
              fill     = "#f58f3b") +

  # Custom Labels;
  labs(title = "",
       subtitle = "",
       x = "Complexity",
       y = "Cutoff") +
  theme_bw(base_size = 17.5, base_family = "Times");

acc.plot <- ggplot() +

  # 5-fold CV Accuracy;
  geom_line(aes(x = c(1:11), y = apply(X = test_error$ACCURACY, MARGIN = 1, FUN = mean)),
            col = "#009dd0",
            size = 0.75) +
  geom_line(aes(x = c(1:11), y = apply(X = train_error$ACCURACY, MARGIN = 1, FUN = mean)),
            col = "#f58f3b",
            size = 0.75) +

  geom_ribbon(aes(x = c(1:11),
                  ymin = apply(X = test_error$ACCURACY, MARGIN = 1, FUN = min),
                  ymax = apply(X = test_error$ACCURACY, MARGIN = 1, FUN = max)),
              alpha    = 0.1,
              linetype = "dashed",
              colour   = "#009dd0",
              size     = 0.75,
              fill     = "#009dd0") +
  geom_ribbon(aes(x = c(1:11),
                  ymin = apply(X = train_error$ACCURACY, MARGIN = 1, FUN = min),
                  ymax = apply(X = train_error$ACCURACY, MARGIN = 1, FUN = max)),
              alpha    = 0.1,
              linetype = "dashed",
              colour   = "#f58f3b",
              size     = 0.75,
              fill     = "#f58f3b") +

  # Custom Labels;
  labs(title = "",
       subtitle = "",
       x = "Complexity",
       y = "Accuracy") +
  theme_bw(base_size = 17.5, base_family = "Times");

sen.plot <- ggplot() +

  # 5-fold CV Sensitivity;
  geom_line(aes(x = c(1:11), y = apply(X = test_error$SENSITIVITY, MARGIN = 1, FUN = mean)),
            col = "#009dd0",
            size = 0.75) +
  geom_line(aes(x = c(1:11), y = apply(X = train_error$SENSITIVITY, MARGIN = 1, FUN = mean)),
            col = "#f58f3b",
            size = 0.75) +
```

```
  geom_ribbon(aes(x = c(1:11),
                  ymin = apply(X = test_error$SENSITIVITY, MARGIN = 1, FUN = min),
                  ymax = apply(X = test_error$SENSITIVITY, MARGIN = 1, FUN = max)),
              alpha    = 0.1,
              linetype = "dashed",
              colour   = "#009dd0",
              size     = 0.75,
              fill     = "#009dd0") +
  geom_ribbon(aes(x = c(1:11),
                  ymin = apply(X = train_error$SENSITIVITY, MARGIN = 1, FUN = min),
                  ymax = apply(X = train_error$SENSITIVITY, MARGIN = 1, FUN = max)),
              alpha    = 0.1,
              linetype = "dashed",
              colour   = "#f58f3b",
              size     = 0.75,
              fill     = "#f58f3b") +

  # Custom Labels;
  labs(title = "",
       subtitle = "",
       x = "Complexity",
       y = "Sensitivity") +
  theme_bw(base_size = 17.5, base_family = "Times");

spe.plot <- ggplot() +

  # 5-fold CV Specificity;
  geom_line(aes(x = c(1:11), y = apply(X = test_error$SPECIFICITY, MARGIN = 1, FUN = mean)),
            col = "#009dd0",
            size = 0.75) +
  geom_line(aes(x = c(1:11), y = apply(X = train_error$SPECIFICITY, MARGIN = 1, FUN = mean)),
            col = "#f58f3b",
            size = 0.75) +

  geom_ribbon(aes(x = c(1:11),
                  ymin = apply(X = test_error$SPECIFICITY, MARGIN = 1, FUN = min),
                  ymax = apply(X = test_error$SPECIFICITY, MARGIN = 1, FUN = max)),
              alpha    = 0.1,
              linetype = "dashed",
              colour   = "#009dd0",
              size     = 0.75,
              fill     = "#009dd0") +
  geom_ribbon(aes(x = c(1:11),
                  ymin = apply(X = train_error$SPECIFICITY, MARGIN = 1, FUN = min),
                  ymax = apply(X = train_error$SPECIFICITY, MARGIN = 1, FUN = max)),
              alpha    = 0.1,
              linetype = "dashed",
              colour   = "#f58f3b",
              size     = 0.75,
              fill     = "#f58f3b") +

  # Custom Labels;
  labs(title = "",
```

```
       subtitle = "",
       x = "Complexity",
       y = "Specificty") +
  theme_bw(base_size = 17.5, base_family = "Times");

grid.arrange(cut.plot, acc.plot, sen.plot, spe.plot, nrow = 2);
}
```

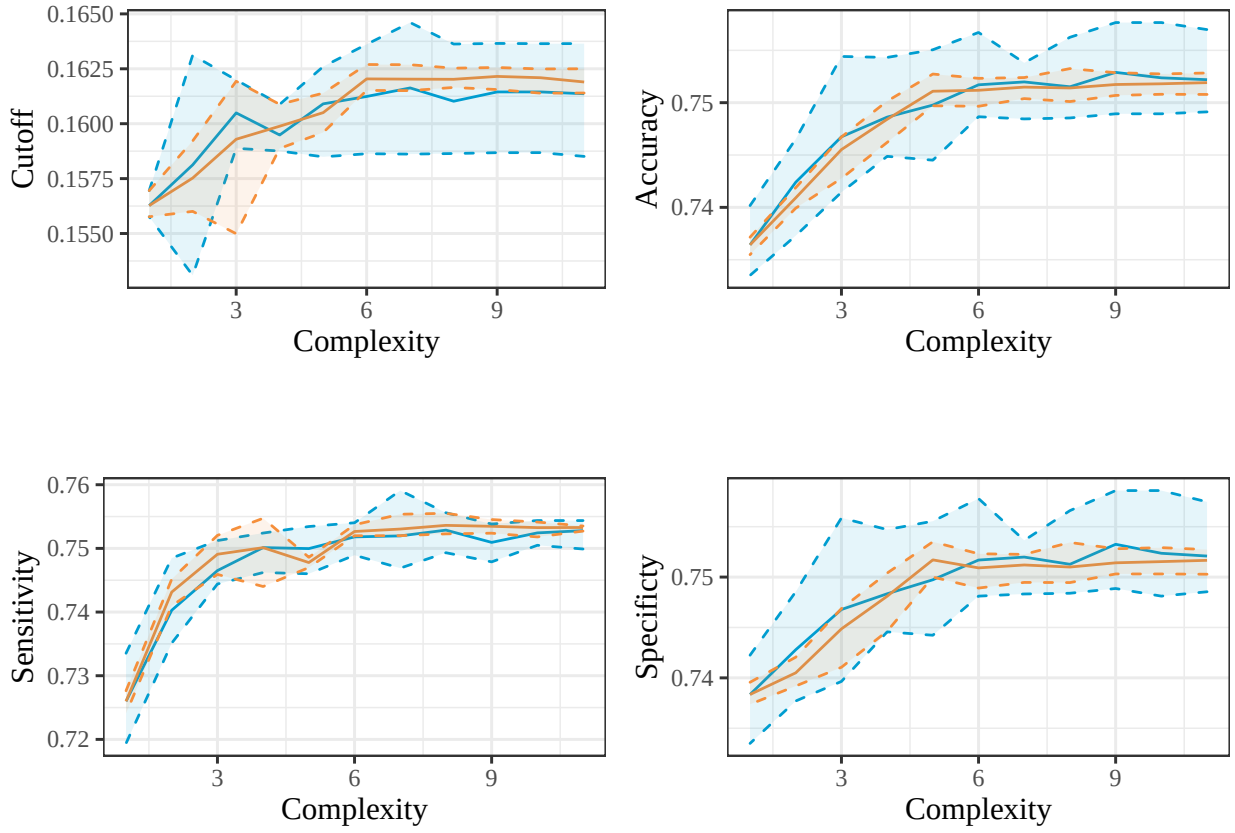

Figure 1: Train (orange) and test (blue) errors for model complexity as average of a 5-fold Cross Validation. The cutoff value is chosen to maximize both accuracy and sensitivity. Dashed lines represent min/max values of the cross validation runs. Top-Left: Cutoff values for model complexity. Top-Left: Accuracy values for model complexity. Bottom-Right: Sensitivity values for model complexity. Top-Right: Specificity values for model complexity.

LOGISTIC REGRESSION SUMMARY;

	Estimates	Standard Error	2.5 (%)	97.5 (%)
(Intercept)	-5.9283660	0.0348357	-5.9968496	-5.8602833
SEXOMale	0.5977156	0.0160538	0.5662831	0.6292166
EDAD	0.0683840	0.0005712	0.0672666	0.0695056
DIABETESYes	0.5262683	0.0178294	0.4912967	0.5611887
EPOCYes	0.1513203	0.0425994	0.0676822	0.2346836
INMUSUPRYes	0.4299422	0.0532283	0.3251884	0.5338675
HIPERTENSIONYes	0.2280410	0.0178166	0.1930922	0.2629341
OBESIDADYes	0.3490336	0.0181811	0.3133573	0.3846291
RENAL_CRONICAYes	0.8865093	0.0398067	0.8084539	0.9645048

LOGISTIC REGRESSION RESULTS;

	Accuracy	Accuracy (SE)	Sensitivity	Sensitivity (SE)	Specificity	Specificity (SE)
Train	0.7514053	0.0013720	0.7536154	0.0013112	0.7509969	0.0018114
Test	0.7515183	0.0030176	0.7528687	0.0025328	0.7512694	0.0033521