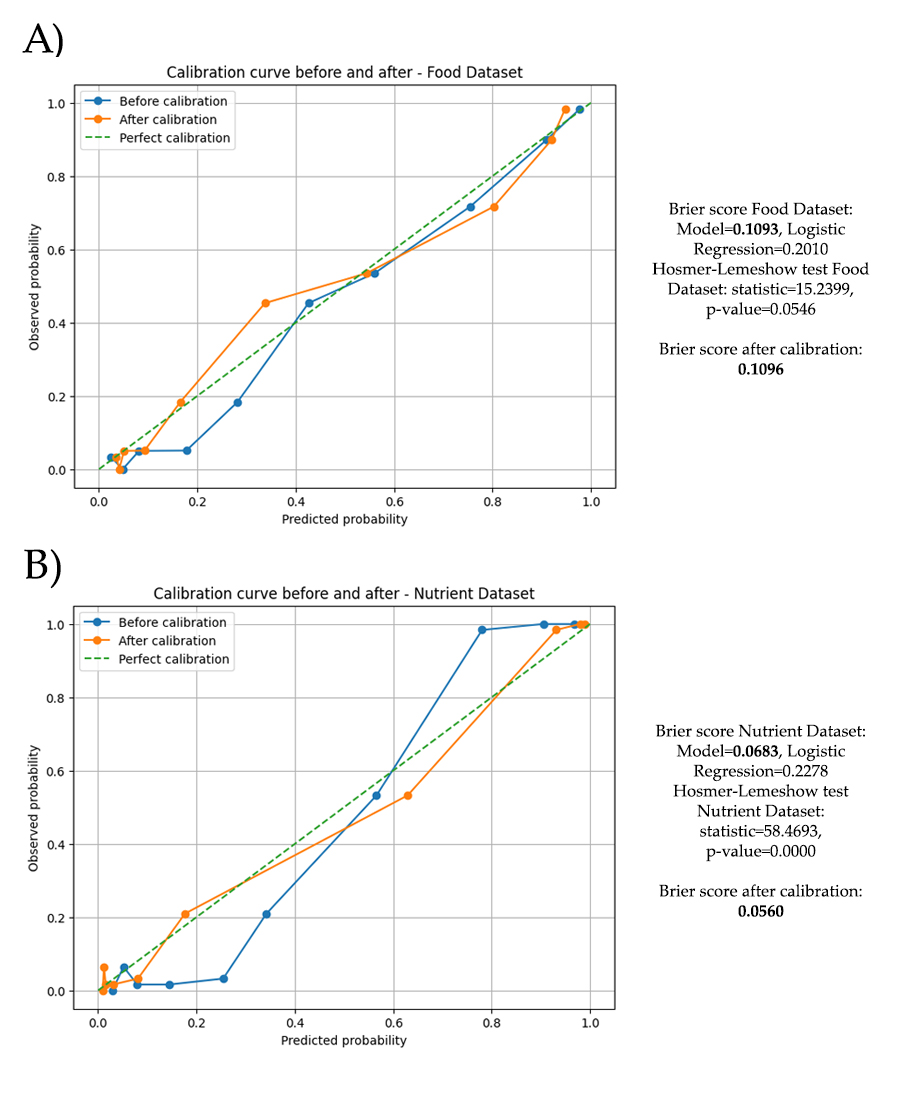
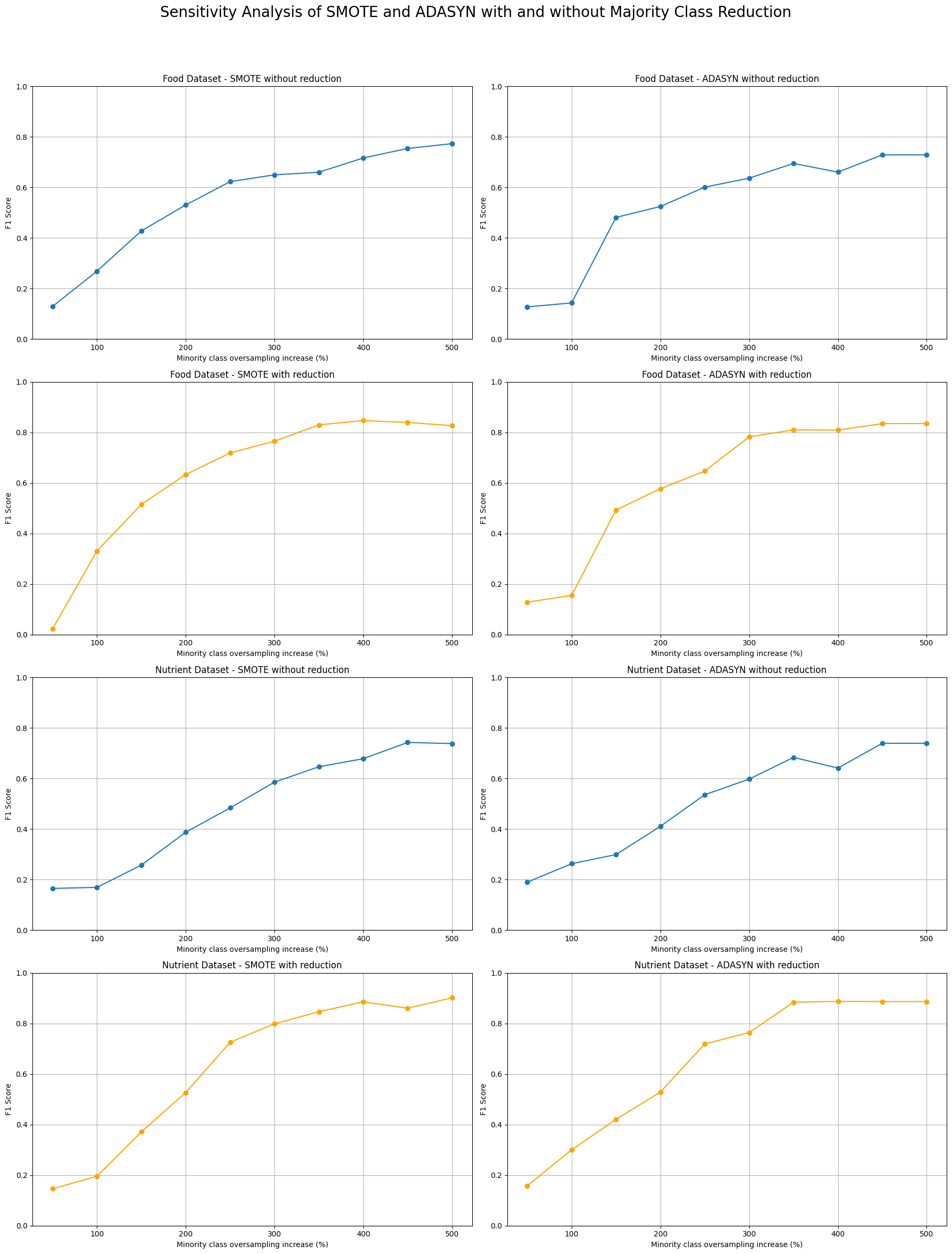
**Supplementary Figure S1.** Calibration curves before and after applying sigmoid calibration to the ensemble model for (A) the Food Dataset and (B) the Nutrient Dataset. In the Food Dataset, calibration did not improve model reliability: the Brier score slightly worsened (from 0.1093 to 0.1096), indicating that post-processing was unnecessary. In contrast, the Nutrient Dataset showed clear gains: the Brier score improved from 0.0683 to 0.0560 and alignment with the perfect calibration line was enhanced, supporting the use of sigmoid calibration under this scenario.



**Supplementary Figure S2.** Sensitivity analysis of oversampling strategies (SMOTE and ADASYN) applied to the Food and Nutrient datasets, with and without majority class reduction. Each panel shows the effect of increasing the minority class oversampling ratio (x-axis) on model performance (F1-score, y-axis). Overall, SMOTE combined with majority class reduction achieved the highest F1-scores, yet the results also highlight that performance improvements can emerge under different resampling scenarios, underscoring the importance of sensitivity analyses when addressing class imbalance.

**Supplementary Table S1.** Sensitivity analysis of SHAP food group interpretations across resampling scenarios. Average feature ranks and directional trends are presented for the main food categories associated with T2DM risk. Fruits, sweets, oils & fats, and red meat consistently emerged as top predictors across scenarios, maintaining stable directional effects. Other groups such as vegetables, tubers, proteins, grains, and white meat showed greater variability in rank positions but preserved consistent directional trends, underscoring the robustness of the core dietary signals while highlighting the influence of class balancing strategies on model interpretation.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Feature** | **Baseline (SMOTE 400%)Avg Rank / Trend** | **SMOTE 250% (50%)Avg Rank / Trend** | **ADASYN 250% (50%)Avg Rank / Trend** | **SMOTE 300% (0%)Avg Rank / Trend** |
| **Fruits** | 4.93 / ↓ risk | 2.15 / ↓ risk | 2.60 / ↓ risk | 1.60 / ↓ risk |
| **Oils & Fats** | 5.96 / ↑ risk | 4.70 / ↑ risk | 4.10 / ↑ risk | 4.55 / ↑ risk |
| **Sweets** | 5.15 / ↑ risk | 2.60 / ↑ risk | 3.45 / ↑ risk | 3.70 / ↑ risk |
| **Vegetables** | 6.52 / ↓ risk | 7.85 / ↓ risk | 8.45 / ↓ risk | 8.10 / ↓ risk |
| **Red Meat** | 4.31 / ↓ risk | 3.30 / ↓ risk | 2.30 / ↓ risk | 2.25 / ↓ risk |
| **Tubers** | 5.23 / ↓ risk | 5.05 / ↓ risk | 5.40 / ↓ risk | 6.55 / ↓ risk |
| **Proteins** | 6.23 / ↓ risk | 6.94 / ↓ risk | 6.61 / ↓ risk | 6.70 / ↓ risk |
| **Hot Beverages** | 5.63 / ↓ risk | 5.40 / ↓ risk | 5.15 / ↓ risk | 4.50 / ↓ risk |
| **Grains & Cereals** | 6.05 / ↓ risk | 7.55 / ↓ risk | 8.06 / ↓ risk | 7.95 / ↓ risk |
| **White Meat** | 6.50 / ↓ risk | 8.67 / ↓ risk | 8.18 / ↓ risk | 8.94 / ↓ risk |

**Supplementary Table S2.** Sensitivity analysis of SHAP nutrient interpretations across different resampling scenarios. The table shows average feature ranks and directional trends for the top predictors of T2DM risk in the Nutrient Dataset. Core nutrients such as zinc, vitamin A, and crude fiber remained consistently influential across scenarios, while vitamin B6 and manganese exhibited changes in directionality, highlighting the importance of sensitivity checks when interpreting model explanations under different oversampling and class reduction settings.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Feature** | **Baseline (SMOTE 400%)Avg Rank / Trend** | **SMOTE 250% (50%)Avg Rank / Trend** | **ADASYN 250% (50%)Avg Rank / Trend** | **SMOTE 500% (70%)Avg Rank / Trend** |
| **Zinc (mg)** | 4.54 / ↓ risk | 3.20 / ↓ risk | 3.00 / ↓ risk | 3.94 / ↓ risk |
| **Vitamin A (µg RE)** | 4.68 / ↓ risk | 3.75 / ↓ risk | 4.35 / ↓ risk | 4.45 / ↓ risk |
| **Crude Fiber (g)** | 5.04 / ↑ risk | 4.61 / ↑ risk | 3.60 / ↑ risk | 4.35 / ↑ risk |
| **Vitamin B6 (mg)** | 4.13 / ↓ risk | 3.21 / ↑ risk | 2.37 / ↑ risk | 1.55 / ↑ risk |
| **Manganese (mg)** | 5.76 / ↓ risk | 5.60 / ↑ risk | 6.00 / ↑ risk | 4.06 / ↑ risk |
| **Potassium (mg)** | 4.67 / ↓ risk | – | 8.00 / ↓ risk | 8.08 / ↓ risk |
| **Animal Protein (g)** | 5.67 / ↑ risk | 5.31 / ↑ risk | 6.71 / ↑ risk | 7.07 / ↑ risk |
| **Vitamin E (mg α-TE)** | 6.60 / ↓ risk | – | – | – |
| **Concentrated Carbohydrates (g)** | 4.23 / ↑ risk | 6.67 / ↑ risk | 6.00 / ↑ risk | 5.85 / ↑ risk |
| **Folic Acid (µg)** | 5.33 / ↑ risk | 5.88 / ↑ risk | 6.00 / ↑ risk | 6.38 / ↑ risk |