Credit Risk Analysis Report:

Overview: This credit risk analysis aims to evaluate the performance of two machine learning models, the original logistic regression model and an enhanced version using the Random Oversampler. The purpose is to assess the models' effectiveness in predicting healthy (0) and high-risk (1) loans.

Original Logistic Regression Model:

- **Accuracy:** 99%
- **Precision (0 - Healthy Loan):** 100%
- **Precision (1 - High-Risk Loan):** 87%
- **Recall (0 - Healthy Loan):** 100%
- **Recall (1 - High-Risk Loan):** 89%
- **F1-Score (0 - Healthy Loan):** 100%
- **F1-Score (1 - High-Risk Loan):** 88%
- **Confusion Matrix:**
  - True Positives (0): 18679
  - False Positives (0): 80
  - False Negatives (1): 67
  - True Positives (1): 558
- **Balanced Accuracy:** 94%

Random Oversampler Model:

- **Accuracy:** 100%
- **Precision (0 - Healthy Loan):** 100%
- **Precision (1 - High-Risk Loan):** 87%
- **Recall (0 - Healthy Loan):** 100%
- **Recall (1 - High-Risk Loan):** 100%
- **F1-Score (0 - Healthy Loan):** 100%

- **F1-Score (1 - High-Risk Loan):** 93%
- **Confusion Matrix:**
  - True Positives (0): 18668
  - False Positives (0): 91
  - False Negatives (1): 2
  - True Positives (1): 623
- **Balanced Accuracy:** 99.6%

Summary:

- Both models demonstrate exceptional accuracy, with the Random Oversampler slightly outperforming the original logistic regression model.
- Precision, recall, and F1-scores are consistently high for both healthy and high-risk loans in both models.
- The Random Oversampler model shows improved performance in predicting high-risk loans, achieving perfect recall and a higher F1-score.
- Confusion matrices indicate minimal misclassifications for both models.

Recommendation: The Random Oversampler model is highly recommended for use by the company due to its superior performance in predicting high-risk loans, as evidenced by its perfect recall and elevated F1-score. The model's balanced accuracy of 99.6% signifies its robustness in handling imbalanced data. The original logistic regression model remains a solid choice, but the enhanced model provides an even more reliable tool for credit risk assessment.