LIST OF NOTATIONS

Symbols	Definition
X_i	Vector of features in the minority class
$\hat{X}_i$	The k-nearest neighbor for X_i
δ	Random number between 0 and 1
Q	The set of queries that are used to compute the attention scores
K	The set of keys that the queries will be compared against to compute the attention scores
V	The set of values that are aggregated to produce the output
d_k	Dimensionality of key vectors
$Softmax$	Function converts the dot product scores into a probability distribution, where each score is transformed into a value between 0 and 1, and the sum of all values is 1
$head_h$	The individual attention and the set of queries that are used to compute the attention scores
$Concat()$	Concatenation of all the individual attention heads
W^O	Output weight matrix
TP	The number of positive samples correctly predicted as positive
TN	The number of negative samples correctly predicted as negative
FN	The number of positive samples incorrectly predicted as negative
FP	The number of negative samples incorrectly predicted as positive