Fold	X_train	y_train	X_val	y_val
1	[[5, 6], [7, 8], [9, 10], [11, 12], [13, 14], [15, 16], [17, 18], [19, 20]]	[0, 1, 0, 1, 0, 1, 0, 1]	[[1, 2], [3, 4]]	[0, 1]
2	[[1, 2], [3, 4], [9, 10], [11, 12], [13, 14], [15, 16], [17, 18], [19, 20]]	[0, 1, 0, 1, 0, 1, 0, 1]	[[5, 6], [7, 8]]	[0, 1]
3	[[1, 2], [3, 4], [5, 6], [7, 8], [13, 14], [15, 16], [17, 18], [19, 20]]	[0, 1, 0, 1, 0, 1, 0, 1]	[[9, 10], [11, 12]]	[0, 1]
4	[[1, 2], [3, 4], [5, 6], [7, 8], [9, 10], [11, 12], [17, 18], [19, 20]]	[0, 1, 0, 1, 0, 1, 0, 1]	[[13, 14], [15, 16]]	[0, 1]
5	[[1, 2], [3, 4], [5, 6], [7, 8], [9, 10], [11, 12], [13, 14], [15, 16]]	[0, 1, 0, 1, 0, 1, 0, 1]	[[17, 18], [19, 20]]	[0, 1]

Figure 1: result of k-fold cross validation.

0.1 Cross-validation

Cross-validation is a technique used in machine learning to evaluate the performance of a model. It involves dividing the dataset into multiple subsets, or "folds", and using one fold as the validation set while the rest of the folds are used to train the model. The process is repeated multiple times, with each fold being used as the validation set in turn, and the performance metrics are averaged across all the folds to get an estimate of the model's overall performance.

Table 1: Sample Data

X1	X2	y
1	2	0
3	4	1
5	6	0
7	8	1
9	10	0
11	12	1
13	14	0
15	16	1
17	18	0
19	20	1

In this example, X1 and X2 are independent variables, and Y is the dependent variable.