Columns with over 10% missing values are removed and in the columns with missing values weighted KNN imputation is performed, since it has been found to be effective in microarray data with 10% missing data [1, 2].

1. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, et al. Missing value estimation methods for DNA microarrays. Bioinformatics. 2001;17(6):520-5.

2. Scheel I, Aldrin M, Glad IK, Sørum R, Lyng H, Frigessi A. The influence of missing value imputation on detection of differentially expressed genes from microarray data. Bioinformatics. 2005;21(23):4272-9.