1. Some of the datalabels are missing, its NaN. Compute the medium value of that feature of the whole dataset and pad it with this value.
2. Using region level features first, its just a wrapped everything in that region, while ROI-level of features is a selected square region of nuclei.
3. Do not drop the first three columns of data, instead, binarize the Tumor and TIL identifier as a feature.
4. Some of the feature columns have a extremely large values, normalize them
5. Build a attention network according to the drawing.
6. Using labels (BCSS), DT and DY column