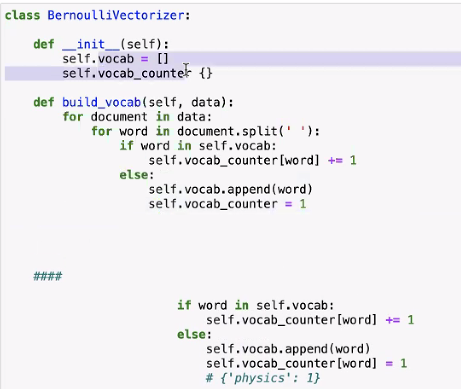
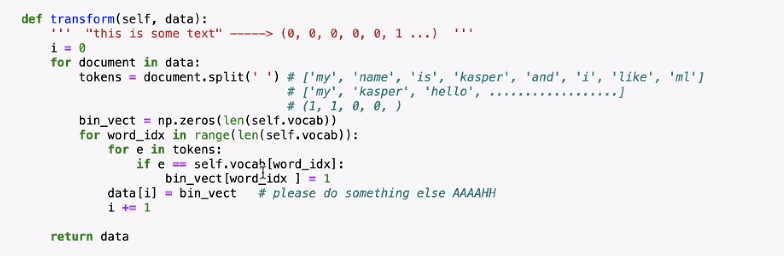




Keep useful 5000 words

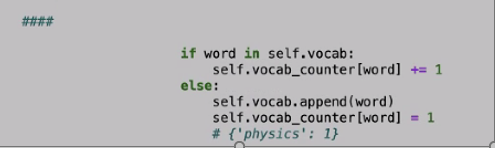
Sort and keep max in the dict.



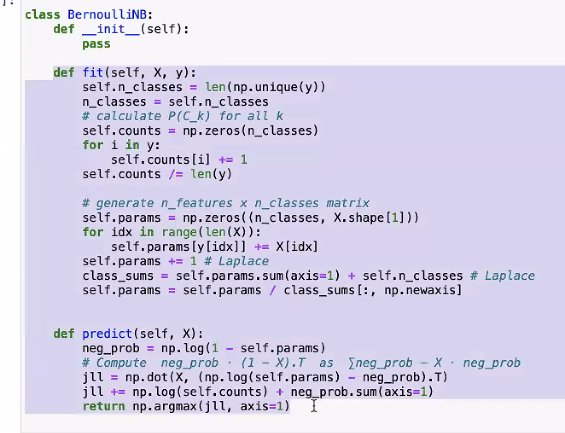


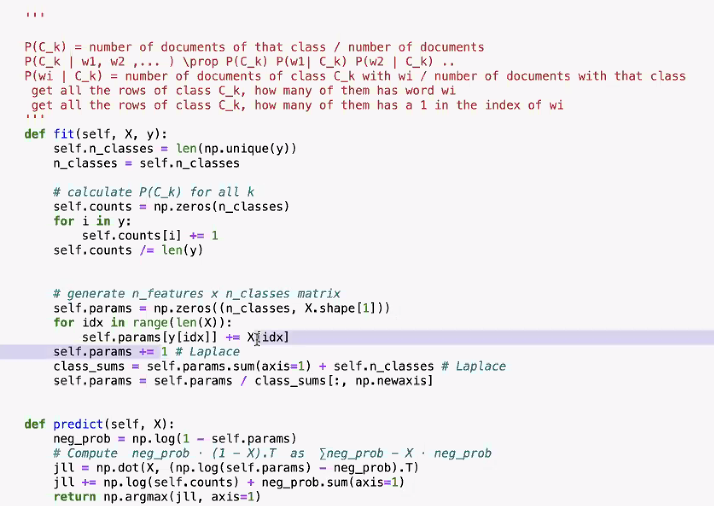


Like one hot encoding.









Use alpha parameter(ie # Laplace)