Preprocessing Raw Data for use in Al Models

