This thesis presents different data mining/machine learning techniques to analyze the vulnerable road users' (i.e., pedestrian and bicycle) crashes by developing crash prediction models at the macroscopic level. In this study, we developed data mining approach (i.e., decision tree regression (DTR) models) for both pedestrian and bicycle crash counts. To the best of the author's knowledge, this is the first application of DTR models in the growing traffic safety literature at the macro level. The empirical analysis is based on the Statewide Traffic Analysis Zones (STAZ) level crash count data for both pedestrians and bicyclists from the state of Florida for the years of 2010 to 2012. The model results highlight the most significant predictor variables for pedestrian and bicycle crash counts in terms of three broad categories: traffic, roadway, and socio-demographic characteristics. Furthermore, spatial predictor variables of neighboring STAZs were utilized along with the targeted STAZ variables in order to improve the prediction accuracy of both DTR models. The DTR model considering spatial predictor variables (spatial DTR model) were compared without considering the spatial predictor variables (aspatial DTR model) and the models comparison results clearly uncovered that the spatial DTR model is superior model compared to aspatial DTR model in terms of prediction accuracy. Finally, this study contributed to the safety literature by applying three ensemble techniques (Bagging, Random Forest, and Boosting) in order to improve the prediction accuracy of weak learner (DTR models) for macro-level crash counts. The model's estimation results revealed that all the ensemble techniques performed better than the DTR model and the gradient boosting technique outperformed other competing ensemble techniques in macro-level crash prediction models.