**INTRODUCTION**

Spatiotemporal data related to the public security have been growing at an exponential rate during the recent years. However, not all data have been effectively used to tackle real-world problems. In order to facilitate crime prevention, several scholars have developed models to predict crime .Most used historical crime data alone to calibrate the predictive models. The research on crime prediction currently focuses on two major aspects: crime risk area prediction and crime hotspot prediction .The crime risk area prediction, based on the relevant influencing factors of criminal activities, refers to the correlation between criminal activities and physical environment, which both derived from the ‘‘routine activity theory’’. Traditional crime risk estimation methods usually detect crime hotspots from the historical distribution of crime cases, and assume that the pattern will persist in the following time periods .For example, considering the proximity of crime places and the aggregation of crime elements, the terrain risk model tends to use crime-related environmental factors and crime history data, and is relatively effective for long-term, stable crime hotspot prediction . Many studies have carried out empirical research on crime prediction in different time periods, combining demographic and economic statistics data, land use data, mobile phone data and crime history data. Crime hotspot prediction aims to predict the likely location of future crime events and hotspots where the future events would concentrate . A commonly used method is kernel density estimation . A model that considers temporal or spatial autocorrelations of past events performs better than those that fail to account for the autocorrelation . Recently machine learning algorithms have gained popularity. The most popular methods include K-Nearest Neighbor(KNN), random forest algorithm, support vector machine (SVM), neural network

and Bayesian model etc. [6]. Some compared the linear methods of crime trend prediction [14], some compared Bayesian model and BP neural network [15], [16], and others compared the spatiotemporal kernel density method with the random forest method in different periods of crime prediction [12]. Among these algorithms, KNN is an efficient supervised learning method algorithm [17], [18]. SVM is a popular machine learning model because it can not only implement classification and regression tasks, but also detect outliers [4], [19]. Random forest algorithm has been proven to have strong non-linear relational data processing ability and high prediction accuracy in multiple fields [20]–[23]. Naive Bayes (NB) is a classical classification algorithm, which has only a few parameters and it is not sensitive to missing data [15], [24]. Convolutional neural networks (CNN) has strong expansibility, and can enhance its expression ability with a very deep layer to deal with more complex classification problems [25], [26]. Long Short-Term Memory (LSTM) neural network extracts time-series features from features, and has a significant effect on processing data with strong time series trends [27]–[29]. This paper will focus on the comparison of the above six machine learning algorithms, and recommend the best performing one to demonstrate the predictive power with and without the use of covariates.