**EXISTING SYSTEM**

Literature has abundant research for Non-Technical Loss (NTL) in electricity fraud detection, but rare researches have been conducted for the water consumption sector. Water supplying companies incur significant losses due to fraud operations in water consumption. The customers who tamper their water meter readings to avoid or reduce billing amount is called a fraud customer. In practice, there are two types of water loss: the first is called technical loss (TL) which is related to problems in the production system, the transmission of water through the network (i.e., leakage), and the network washout. The second type is called the non-technical loss (NTL) which is the amount of delivered water to customers but not billed, resulting in loss of revenue. To address these challenges, Jordan ministry of water and irrigation as in many other countries is striving, through the adoption of a long-term plan, to improve services provided to citizens through restructuring and rehabilitation of networks, reducing the non-revenue water rates, providing new sources and maximizing the efficient use of available sources. At the same time, the Ministry continues its efforts to regulate the water usage and to detect the loss of supplied water.

**PROPOSED SYSTEM**

This paper focuses on customer’s historical data which are selected from the YWC billing system. The main objective of this work is to use some well-known data mining techniques named Support Vector Machines (SVM) and K-Nearest Neighbor (KNN) to build a suitable model to detect suspicious fraudulent customers, depending on their historical water metered consumptions. The CRISP-DM (Cross Industry Standard Process for Data Mining) was adopted to conduct this research. The CRISPDM is an industry standard data mining methodology developed by four Companies; NCR systems engineering, DaimlerChrysler AG, SPSS Inc. and OHRA. The CRISP-DM model consists of business understanding, data understanding, data preparation, model building, model evaluation and model deployment. To extract the fraud customers’ profile, a new table is created containing the client's number, the water consumption, and a new attribute for fraud class. This attribute is filled with a value of ‘YES’. Another table for the normal clients is created, and the fraud class attribute is filled with the value “NO”. The two tables are then consolidated into one table containing the customer ID, consumption profile, and fraud class attributes. To filter the data, some preprocessing operations were performed such as Eliminate redundancy, Eliminate customers having zero consumption through the entire period, Eliminate new clients who are not present during the whole targeted period, and Eliminate customers having null consumption values. Filtering the data resulted in a reduced original dataset of the non-fraud customer to 16114 record and the fraud customers to 647 records.