

CHAPTER 5 SVM-UMAP FOR CHILD MALNUTRITION PREDICTION ANALYSIS

*Joynal Abdin, Habibollah Haron,
Suhaila Mohamad Yusuf, and Mohamad Shukor Talib*

5.1 INTRODUCTION

The human body and health depend on nourishment, which is necessary for survival. An inadequate calorie intake, a lack of extra calories, and an unbalanced energy level characterise malnutrition. Lack of vital minerals and vitamins will result in several major health issues. Malnutrition can refer to all forms of inadequate nutrition. Malnutrition is divided into two cats, undernutrition and overnutrition, which are the conditions for a child's growth. When a person is "over-nourished or overweight" due to a high-calorie diet, an unhealthy lifestyle, and a health issue. Numerous diseases, including cancer, diabetes, and heart disease, are caused by it. Children are deemed to be undernourished if they are lacking in critical minerals and proteins. When toddlers have mineral and vitamin deficits, it is known as undernutrition. It is defined as wasted, underweight and stunted.

According to Fanzo et al. (2018) report, 2.6 million children die each year due to malnutrition worldwide, and 45% of children die before the age of five. In addition, it was revealed that 38.3 million children were

overweight or obese, 14.3 million were extremely underweight, and 144 million were stunted. Malnutrition is caused by poverty, an improper diet, chronic illness, illiteracy, and population expansion, and 46% of toddlers in Bangladesh struggle with being underweight. World Health Organization (WHO) reported that in the year 2023, there were 45.4 million wasted children, 149.2 million stunted children, and 38.9 million overweight children.

5.2 RESEARCH IN MALNUTRITION

Malnutrition is one of the main problems that are wholly or partially related to children's deaths worldwide among children aged between zero months to 59 months. Demographic, physical, socioeconomic, health and household factors are categorised as the cause of children's malnutrition. Demographic and health survey (DHS) for Asian and African countries like Bangladesh, Nepal, Sri Lanka, India, Ethiopia, etc. Child malnutrition in Bangladesh is the most malnourished country in the world. In the survey by DHS 2011 in Nepal, the outbreak of children under five years of age were stunted at 41%, underweight at 29%, and wasted at 11%, respectively (Talukder & Ahammed, 2020).

Various machine learning (ML) algorithms have recently been employed in the healthcare industry to identify the problem and its root cause from nutritional data. ML algorithms have been used in medical research to predict diseases. For instance, artificial neural networks (ANN), random forest (RF), and support vector machines (SVM) have been used to denote the status of diseases and predict diseases such as heart disease and diabetes (Reis et al., 2017).

In prior studies, Khare et al. (2017) explored data from the Indian DHS. Thangamani and Sudha (2014) examined data from the Indian family health survey, while Markos (2014) analysed data from the Ethiopian DHS to assess and forecast the health status of Ethiopian children. Processing data with a high dimension increases the execution time required and the complexity of memory consumption. ML cannot perform satisfactorily when dealing with high-dimensional data.

Therefore, dimensionality reduction can be applied to reduce the number of input variables in training data. It is a method for preparing data that is carried out before applying it to ML modelling. Before beginning the training process for a predictive model, the data must be cleaned and scaled (Witten et al., 2016). The number of variables can be reduced by dimensionality reduction, which improves the classification's performance.

Dimensional reduction is made possible using manifolds and projection technologies. It provides a fast-running time and a high level of repeatability, and it may be used to analyse high-dimensional data of any data format (Li et al., 2022). The UMAP dimension reduction is used for supervised, unsupervised, and semi-supervised ML techniques (McInnes & Healy, 2018).

5.2.1 Techniques for Predicting Malnutrition

Various ML algorithms have recently been employed in the healthcare industry to identify the problem and its root cause from nutritional data. Currently, the malnutrition status can be predicted using an algorithm for ML to analyse malnutrition data. ML algorithms have been used in medical research to predict diseases. For instance, ANN, RF, and SVM have been used to denote the status of diseases and predict diseases such as heart disease and diabetes.

Previously, Khare et al. (2017) investigated Indian DHS data to analyse Indian children and calculated HAZ, WHZ, WAZ, and BMI from the IDHS dataset. Research by (Thangamani & Sudha, 2014) studied Indian family health survey data to predict undernutrition based on children's food consumption and decision trees. ANN algorithms were applied to that study. Markos (2014) studied the Ethiopian Demographic and Health Survey data (EDHS) for analysing and predicting Ethiopian children's health status. They applied PART and Naïve Bayes (NB) algorithms to EDHS data for predicting the malnourished status of toddlers and calculated WHZ, WAZ and HAZ for children.