


## Article

# Adapting Fleming-Type Learning Style Classifications to Deaf Student Behavior

Tidarat Luangrungruang and Urachart Kokaew \* 

Applied Intelligence and Data Analytics Laboratory, College of Computing, Khon Kaen University, Khon Kaen 40002, Thailand; tidarat.l@kkumail.com

\* Correspondence: urachart@kku.ac.th

**Abstract:** This study presents the development of a novel integrated data fusion and assimilation technique to classify learning experiences and patterns among deaf students using Fleming’s model together with Thai Sign Language. Data were collected from students with hearing disabilities (Grades 7–9) studying at special schools in Khon Kaen and Udon Thani, Thailand. This research used six classification algorithms with data being resynthesized and improved via the application of feature selection, and the imbalanced data corrected using the synthetic minority oversampling technique. The collection of data from deaf students was evaluated using a 10-fold validation. This revealed that the multi-layer perceptron algorithm yields the highest accuracy. These research results are intended for application in further studies involving imbalanced data problems.

**Keywords:** classification; deaf students; feature selection; imbalanced data; learning style



**Citation:** Luangrungruang, T.; Kokaew, U. Adapting Fleming-Type Learning Style Classifications to Deaf Student Behavior. *Sustainability* **2022**, *14*, 4799. <https://doi.org/10.3390/su14084799>

Academic Editor: Rosabel Roig-Vila

Received: 14 March 2022

Accepted: 14 April 2022

Published: 16 April 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Education is a vital social catalyst for improving quality of life, and everyone (including those with physical, emotional, or intellectual disabilities) should expect equal access to education. The Royal Thai Government, in support of equal rights and educational opportunities for its people, established the National Education Act, B.E. 2542 (1999). That Act states: “People with physical, mental, intellectual, emotional, social, communication, and learning deficiencies; those with physical disabilities; the crippled; those unable to support themselves; or those destitute or disadvantaged; will have the right and opportunities to receive basic education specifically provided” [1].

The educational provision of special classes for deaf individuals needs to be undertaken carefully. As a starting point, the nature and style of learning should be determined to allow for the enhancement of student skills and the development of appropriate instruction. Therefore, teachers must understand the diversity and learning styles of special learners so as to utilize the right teaching tools and help students learn effectively. For this purpose, digital academic data can be obtained and converted using data mining techniques that elicit significant knowledge from the data [2]. Once mined, the data can be classified into learning style types that improve learning methods. The application of data mining to educational data is referred to as educational data mining (EDM) [3,4]. Some technological approaches, such as e-learning, computer-assisted instruction, the World Wide Web, and multimedia, have also been adopted for classroom use [5,6]. Technology is applied to create learning media for deaf students that are informed by and meet their educational needs.

This study aims to analyze the main factors affecting learning styles by comparing learning type classifications via data synthesis and resolving data imbalance. These efforts resulted in the identification of factors influencing learning styles and the techniques that yielded the highest accuracy from each classification. When learning styles matched needs, the appropriate instructional method was provided to that individual student [7], allowing for improved academic achievement.

## 2. Related Work

### 2.1. The Algorithm Used for Data Analysis

In the past few years, it has been easier to gather learning process data via computer-based learning environments. This has led to a growing interest in the field of learning analytics. This involves leveraging learning process information to better understand and ultimately improve teaching and learning [8]. Oranuch Pantho and Monchai Tiantong [9] classified the VARK learning styles using Decision Tree C4.5, in which student learning styles were obtained via a questionnaire. The data included gender, age, length of study, GPA, and educational background. Their study concluded that Decision Tree C4.5 effectively classified the students' VARK learning styles. Worapat Paireekreng and Takorn Prexawanprasut [10] determined classifications using four different algorithms: decision tree; naïve Bayes; artificial neural network (ANN); and support vector machine (SVM). The indicators measuring model efficacy were dependent on the accuracy of the classification, machine learning, and learning style determination. The proposed mechanism improved the classifications using k-nearest neighbor (K-NN) with the genetic algorithm (GA), which was carried out in an open learning management system [11]. According to David Kolb, learning style classifications and comparisons of classification efficacy were implemented using the decision tree, naïve Bayes tree (NBT), and naïve Bayes algorithms. Tenfold cross-validation was adopted to build and test the model and to analyze the data [12]. Several classifiers, such as Bayes (naïve Bayes and Bayes net), tree (J48, NB Tree, random tree, and CART), and rules (conjunctive rule, decision table, and DTNB) tested the efficiency of the learning style classifications, involving the preferences and behavior of students using e-learning. The resulting features were analyzed and classified into learning styles based on the Felder–Silverman learning style model. Tenfold cross-validation was used to evaluate the classifiers [13]. Decision trees, Bayesian family classifiers, and neural networks were the methods used in this study. Decision trees are inherently interpretable and provide a visual representation [14], and the Bayesian family classifiers perform with superior efficiency. In addition, these methods allow the model to be constructed more quickly. Thus, the Bayes net and naïve Bayes classifiers were chosen for this investigation [15]. The neural network is a basis for deep learning. A large amount of generated data are available in a deep learning blend of deep learning-based neural networks [16].

In summary, based on the review of the previously mentioned literature, the methods chosen to analyze data in this study were the decision tree, random forest, Bayesian network, naïve Bayes, multi-layer perceptron, and k-nearest neighbor algorithms. The algorithm chosen for this research is the most popular for data classification analysis and yields highly accurate results.

### 2.2. Analysis of Learning Factors

Much educational research has explored the extent to which learner variables can predict learning outcomes or future learning behaviors [17]. The investigation of learning styles has significantly enhanced our understanding of how adult learners acquire knowledge via an online curriculum [18]. Electronic catalogs have been designed and built based on the purchasing habits of consumers [19], utilizing VARK to analyze similar information such as gender, age, race, and experience in using both a computer and the internet. Several educational studies have adopted various factors to evaluate the VARK learning styles of first-year students in order to meet various academic demands [20]. Any educational system must be able to support different content and instructional media for all potential learners, thereby providing the most effective learning possible [10]. Such systems must be able to check the correlation between learning style and learning efficacy via problem-solving games and effective teaching strategies [21]. The analysis of gender, age, and education level assists in the determination of learning style [10,20,21]. Questionnaires were provided to elicit the level of satisfaction with VARK learning styles, allowing learners to overcome weaknesses, select their preferred learning style, and improve their academic outcomes [22]. Additionally, other factors such as GPA [10,22], fields of study [21,23], and

the student's hometown [20,22] were studied for their effects on selected learning styles. All of these factors affect learning, and understanding and adapting to them results in better learning. These factors were therefore selected for analysis in this research.

### 2.3. Learning Style

Inconsistencies between teaching approaches and student learning styles may result in negative consequences, such as a lack of attention and interaction, and an inability to pass tests and classes, which can eventually result in the student leaving the school. Therefore, educators need to balance their teaching methods with a learning model that responds to a student's needs [24]. Learning style is the personal style that an individual uses in a learning task [25].

Individual learners have different learning styles. They evaluate data and information through their perceptions in different ways. A study of engineering students [26] identified four types of learners: the diverger, the assimilator, the converger, and the accommodator [27,28]. These classifications were then expanded into the investigation of other learning styles in fields such as science. These learning styles were divided into five dimensions with two opposite connecting sides: (1) sensing/intuitive; (2) visual/verbal; (3) inductive/deductive; (4) active/reflective; and (5) sequential/global [29]. Fleming [30] further proposed a sensory model developed from Eicher's work [31], called VARK: visual (V), aural (A), read/write (R), and kinesthetic (K) learning styles. The VARK model is defined as an individual way to collect, manage, and organize ideas and data senses, and it is included in pedagogy due to its relationship with data reception and contribution [32]. Fleming's VARK learning styles were based on four sensory perceptions: visual, learning through seeing; aural, learning through listening and discussing, including music and lectures; learning through reading and writing; and kinesthetic, learning through body movement, expressing feelings, and touching [30,33,34]. The VARK learning styles have been adopted and applied in other fields, such as business, programming, nursing, online learning, and physiology, and have also been adopted to check the acceptance of different educational technologies by learners [19,35–40].

### 2.4. Learning for the Deaf

The most common form of communication is verbal expression and normal language [41]. People with a hearing disability can find it difficult to interact verbally with their peers and must be attentive to the sounds around them [42]. Around the world, many deaf and hard-of-hearing people benefit from offline subtitles (e.g., for pre-recorded television programming) or real-time captioning services (e.g., in classrooms, meetings, and at live events) [43]. Although deaf people are unable to hear, their other senses, such as sight, assist them in 'listening' to others. Facial expressions during speech assist the listening process, in particular lip reading [44]. Where hearing individuals can process errors and make corrections to what they hear, lip readers hone their skills for the same purpose. Developing the ability to understand someone speaking using lip reading, as well as learning to 'sign' language, is particularly difficult [45], and, obviously, sign language is only useful for communication between knowledgeable participants.

Sign language within the deaf community consists of body movements, facial expressions, hand positions, and body posture [46–48], which are expressed in three dimensions [49]. They range from individual ad hoc contexts to official, national sign languages [50]. Thai Sign Language is widely broadcast on several television networks offering content translation, and this exposure further serves to raise the profile and awareness of sign language [51]. Deaf children communicate using sign language and by reading lips.

Because reading plays a significant role in learning [52], deaf individuals must memorize characters rather than sounds. In writing, a deaf student typically writes simple sentences with a fixed structure and no transition and is unable to connect complex sentences [53]. In addition, mistakes are likely [54]. For these reasons, specific learning styles

for deaf students are needed. It is very important that teaching and learning is based on the preferences and unique needs of deaf students.

### 3. Methodology

#### 3.1. Data Analysis Techniques

Data analysis techniques used to search for patterns, hidden relationships, or existing rules in large data are referred to as data mining [55]. Identified patterns may point to particular meanings that can be utilized for certain purposes [56–58]. In this study, the process involved finding patterns and relationships hidden in the deaf student data set. The integrated analysis applied in this study was based on basic data relating to deaf students, learning style information, and algorithms. The algorithm for classifying learning styles helped to build models in response to receiving new information. The various algorithms, including decision tree, random forest, Bayesian network, naïve Bayes, k-nearest neighbor (KNN), and multi-layer perceptron, used to classify the data [59] are described in the following subsections.

##### 3.1.1. Decision Tree

Developed from the ID3 algorithm and named after a hierarchical model visually formed in the shape of a tree, decision tree (C4.5) has become the standardized comparative tool for learning algorithms [2]. It represents a learning process, in which data are classified based on their attributes [60]. The algorithm splits observations into multiple branches, also referred to as subsets, based on a decision node with a given criterion [61].

Attributes must be selected as root nodes for this algorithm. A gain criterion is used to select the best attribute. Using information gain reduces the number of classification tests, making the tree less complex. The information gain equation is as follows:

$$(s_1, s_2, \dots, s_n) = - \sum_{i=1}^n \frac{s_i}{s} \log_2 \frac{s_i}{s} \quad (1)$$

where

$s$  refers to the number of data sets (e.g.,  $s$  records);

$n$  refers to the total number of different groups in the data set;

$C_i$  refers to the group of order  $i$  where  $i = 1, \dots, n$ ;

$s_i$  refers to the number of data points that belong to  $s$  in group  $C_i$ .

##### 3.1.2. Random Forest

Random forest, 2001 [62], is built by constructing several models using decision trees and randomized variables. Random forest is an effective and well-established method for generating multiple classifications and regression trees, based on bagging and random subspace [63]. The outcomes of each model are combined, and the most-repeated one is selected and extracted as the outcome, which in turn forms the structure of the tree. The advantages of this method are forecast precision, user-friendliness, overall effectiveness, and the ability to work with unseen information (as it is less over-fitting than other classifications [64]). Its ease of use derives from it requiring only two parameters: the variables in a random set of each node and the number of trees [65].

##### 3.1.3. Bayesian Network

The Bayesian network classification method, developed from the law of Bay [66], is conducted by plotting a probability graph (known as Bayes net) that reduces limitations introduced by naïve Bayes (see Section 3.1.4). Thus, the Bayesian network is used to explain the independence within conditions between variables. To enhance learning effectiveness, previous knowledge should be input into the Bayesian belief network (Equation (3)) in the form of a network structure that includes the probability table and conditions. For example,

$X$  is conditionally independent of  $Y$  (which means the probability of  $X$  does not depend on  $Y$ ) when  $Z$  is known; this is written in the following equation:

$$P(Y, Z) = P(Z) \quad (2)$$

In a Bayesian network, each variable contains a particular probability that may derive from a beginning node or the relationship of more than one node. A probability occurring from more than one variable, referred to as a joint probability, is expressed using the following equation:

$$P(x_1, \dots, x_n) = \prod_{j=1}^n P(\text{Parents}(x_i)) \quad (3)$$

where  $x_i$  refers to the variable considered from *parents* and where  $(x_i)$  refers to the variable considered from the case of direct parents of  $x_i$ .

#### 3.1.4. Naïve Bayes

Naïve Bayes classification uses statistical probability to forecast membership according to Bayes' theorem [2]. Bayes enhances the learning model by adding training sets, which accounts for its popularity and widespread use [67]. The algorithm is uncomplicated, learns quickly, and can manage a variety of features or classes within a variety of cases [68]. However, the outcomes are effective only when the data features selected are independent of each other, as written in Equation (4):

$$P(A) = \frac{P(A|C) \times P(C)}{P(A)} \quad (4)$$

where

$P(C)$  refers to the probability of the incident before incident  $C$  occurs;

$P(A)$  refers to the probability of the incident before data set  $A$ ;

$P(C|A)$  refers to the probability of incident  $C$  when incident  $A$  occurs;

$P(A|C)$  refers to the probability of incident  $A$  when incident  $C$  occurs.

#### 3.1.5. K-Nearest Neighbor

KNN [69] is considered one of the ten most suitable algorithms for data classification due to its simple and effective process [59]. Adopting the "lazy learner" concept, KNN states that classifiers from data sets are not needed; instead, it collects data, waits to reach the target, and then starts its analysis. The principle of KNN is similar to that of data clustering: the distance between the predicted data and all nearby data (number of data =  $k$ ) is measured. This is a widely used fundamental classification [70], in which the predicted outcome is classified within the entire KNN. The distance measurement is conducted in the Euclidean style, where the second root of the attribute differences is squared, as shown below:

$$D_{euclidean} = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + \dots + (X_L - Y_L)^2} \quad (5)$$

where  $X_1$  refers to the first attribute of data point 1 and  $Y_1$  refers to the first attribute of data point 2, in which both data ( $X$  and  $Y$ ) contain  $L$  attributes.

#### 3.1.6. Multi-Layer Perceptron

The multi-layer perceptron, patterned within an artificial neural network (ANN), is a data-processing model first developed in 1943 by McCulloch Pits as a simple ANN of "McCulloch-Pits" neurons [71]. The concept, inspired by the bioelectric network of neurons and synapses in the human brain, processes data using calculations in a network where several subprocessors work together. Its multi-layer structure is effective in solving complex tasks using supervised learning in backpropagation networks. The architecture sends data



from the input layer to the hidden and output layers, where the processing direction of data flows back to fix errors in every hidden layer, thereby improving the operation.

### 3.2. Dividing Data to Test the Efficiency of the Classification Model

Cross-validation [72,73] (using self-consistency testing, split testing, and k-fold cross-validation) plays a significant role in measuring the efficiency of a forecasting model. The data are divided into learning and testing data sets for classification. If cross-validation is not selected carefully at this stage, the classification outcomes may contain errors. Therefore, k-fold cross-validation was chosen for model performance testing, as it is a popular method and provides reliable results [74,75].

K-fold cross-validation has become widely used due to its reliability [76]. Testing a model's efficiency using cross-validation is achieved by dividing data into k groups (where k is any value from 1 to n), in which each group contains the same amount of data. A single group is then selected as the test group, and the remaining groups become learning groups. The test group is circulated in k rounds until all groups are completed and all data classified. The most common method, 10-fold validation, provides positive results but takes some time.

### 3.3. Information Gain

Feature selection is a useful way to reduce feature space dimensions. By developing data collection and machine learning techniques, feature selection plays an important role in data mining and machine learning. Feature selection can not only extract significant impact factors but can also improve accuracy [77].

Feature selection is conducted by calculating and ordering the weights correlated between features and classes using the following equation:

$$\begin{aligned} \text{Information Gain} &= \text{Entropy}(\text{initial}) \\ &- [P(c_1) \times \text{Entropy}(c_1) + P(c_2) \times \text{Entropy}(c_2) + \dots] \end{aligned} \quad (6)$$

where  $\text{Entropy}(c_1) = -P(c_1)p(c_1)$  and  $p(c_1)$  refers to the probability of  $c_1$ .

According to Equation (6), information gain relies on measures of differences and dispersions of data, known as entropy. If data are either significantly different or significantly similar, the entropy will be high. More details of entropy can be found in [78].

### 3.4. Synthetic Minority Oversampling Technique

The synthetic minority oversampling technique (SMOTE) [79] randomizes the determined amount of data from the minority class via KNN. The nearest k is chosen to build synthesized samples within the area where vector properties and the nearest k are calculated to find the distance between vectors. The differences are multiplied by random numbers between 0 and 1, which are added to the feature:

$$N_{\text{point}} = O_{\text{point}} [0, 1] + (\text{Random}[0, 1] \times \text{distance}(x, y, \dots, z)) \quad (7)$$

where

$N_{\text{point}}$  refers to a newly developed data point of the minority class;

$O_{\text{point}}$  refers to a data point of the minority class as a starting point of the distance compared with the neighboring point;

$\text{Random}[0, 1]$  refers to a random number between 0 and 1;

$\text{distance}(x, y, \dots, z)$  refers to the distance between the starting point and the neighboring point from attributes  $x$  and  $y$  to  $z$ .

### 3.5. Index of Item–Objective Congruence

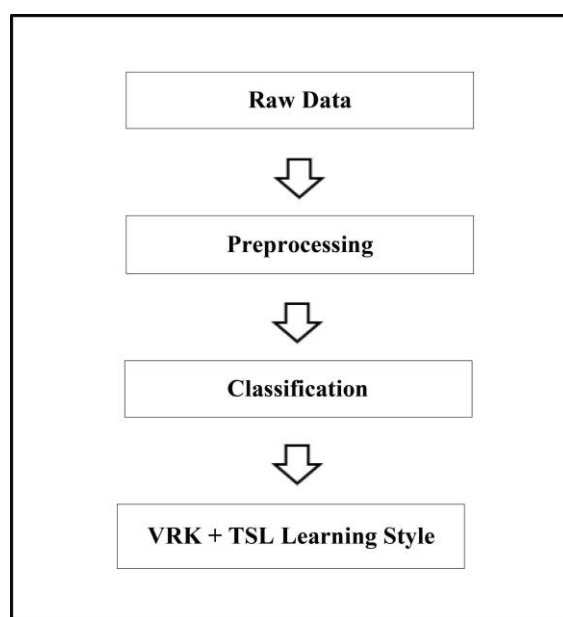
Item–objective congruence (IOC) is an innovative procedure that is used to evaluate content validity, objective items, and questions [80]. Analysis of the index of items evaluates

research tools, with data being collected via questionnaires. The objectives are presented to three to five experts specializing in data assessment, to consider whether the tools are consistent with the objectives. Content validity, language correctness, and question clarity are checked and the IOC determined.

### 3.6. Research Methodology

This research created a model for classifying the learning styles of deaf students who attended Udon Thani and Khon Kaen schools for the deaf and who could communicate in Thai sign language. The objective was to identify a learning style suitable for each student. A limitation was that these were students from the first group of schools to initiate the teaching of deaf students. This was therefore a niche group that might result in a smaller population. Udon Thani and Khon Kaen are special schools for deaf students and were selected because they have the largest number of students in the northeastern region of Thailand. The research was endorsed by the Office of the Khon Kaen University Ethics Committee in Human Research to run from 3 February 2021 to 20 January 2023. The hypothesis was that deaf students were able to determine their appropriate learning style from the relevant factors.

Although Fleming's learning style model has been widely used to classify learning styles, it has never been applied to classify the learning styles of deaf students or to Thai Sign Language (TSL) as a communication language for the deaf. This hybrid learning style for the deaf is VRK + TSL. The research work plan of the VRK + TSL Rule Model is shown in Figure 1.



**VRK + TSL Rule Model**

**Figure 1.** Research work plan.

#### 3.6.1. Work Plan of the VRK + TSL Rule Model

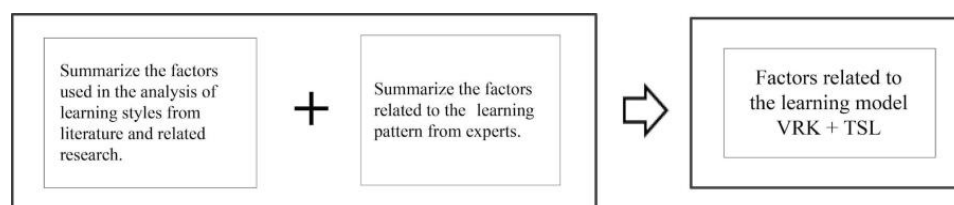
Analysis of the VRK + TSL learning pattern classification was categorized into three parts:

1. The related predictor of VRK + TSL learning (Figure 2).

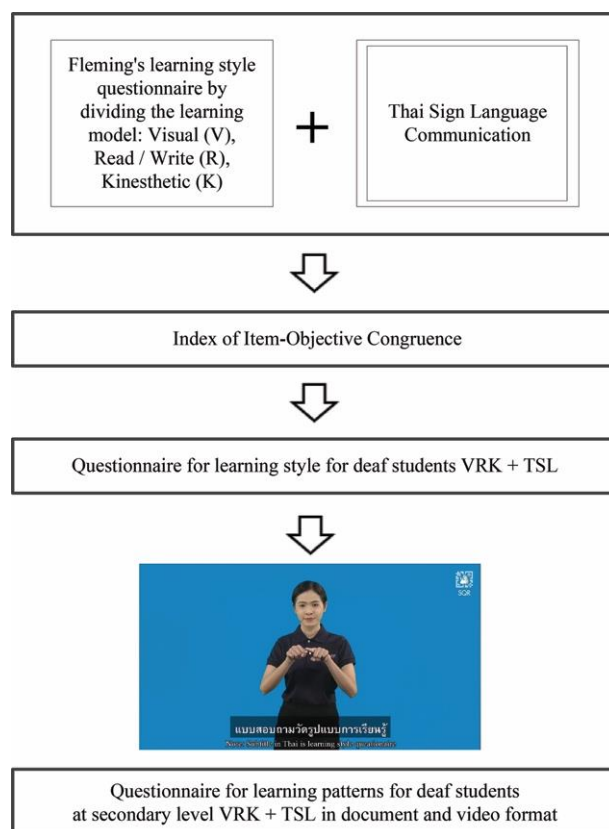
Exact predictors for VRK + TSL learning have not yet been determined. Several literature sources have concluded that “bunches” of data should be employed (see Section 2.2) and we adapted these factors [81] to analyze deaf students. Selection of the most appropriate predictor can be achieved using the following steps:

- A questionnaire is developed to determine the predictor for the VRK + TSL learning style, which investigates the factors that affect learning among the deaf. A Likert scale is then used to evaluate the results [82].
  - The questionnaires are analyzed by five experts.
  - The expert comments are gathered and averaged scores greater than 3.50 are used to construct the appropriate learning pattern.
2. VRK + TSL learning pattern questionnaires for deaf students (Figure 3).

A questionnaire was constructed based on the VRK + TSL pattern combining factors used for analyzing VARK + TSL learning (Step 1), employing the VARK developed by Neil Fleming. The questionnaire consisted of 16 items with four choices representing four aptitudes: visual, read/write, kinesthetic, and Thai Sign Language, adapted to include VRK + TSL content. The time needed to complete the questionnaire was approximately 30 min. The factors affecting learning among deaf students were analyzed by five experts. Data were collected and analyzed to rate the questionnaires using basic statistics. Content validity, appropriateness of language, and question clarity were also reviewed. The questionnaires exhibited significant ratings (0.60–1.00), implying an index of IOC greater than 0.70. The language was then clarified based on the suggestions of the experts, to ensure participant understanding.



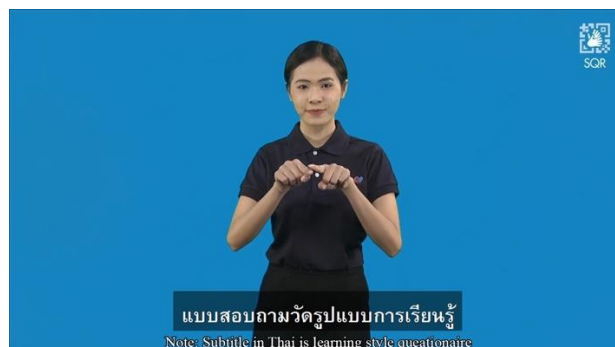
**Figure 2.** Factors related to the VRK + TSL learning model.



**Figure 3.** Creating a questionnaire based on the VRK + TSL learning model.



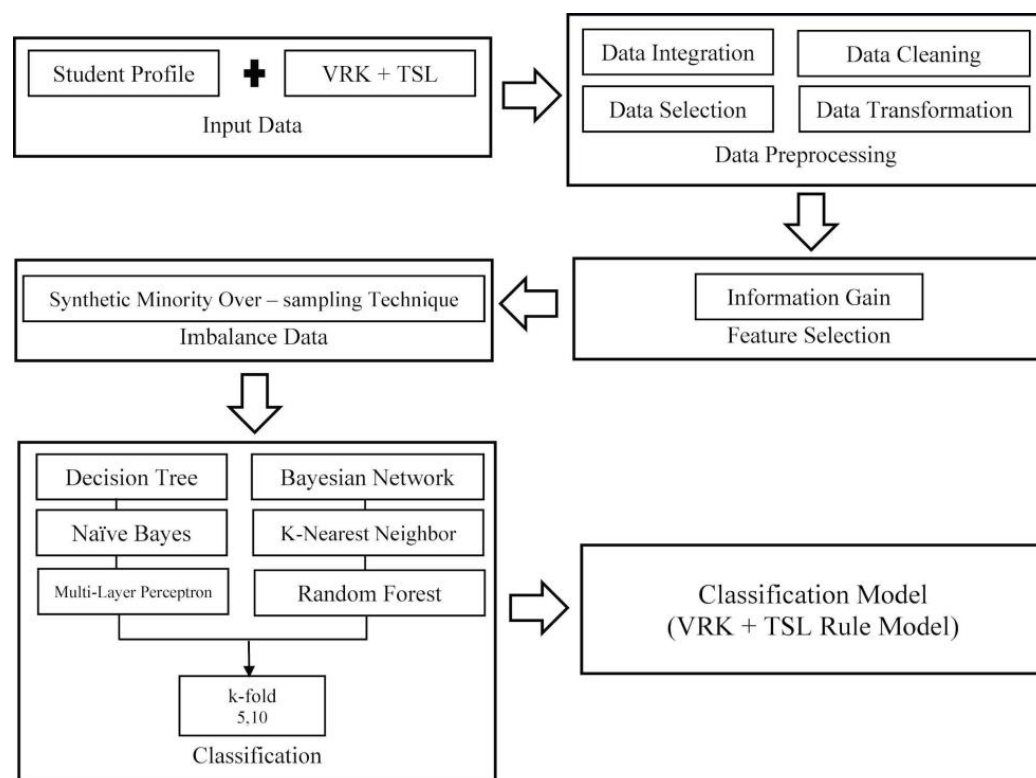
Learning style questionnaires were incorporated into a TSL video together with documents for deaf students in Thai secondary schools. Under the supervision of the National Electronics and Computer Technology Center (NECTEC), a QR code was embedded in the documents to connect with the sign language video, as shown in Figure 4.



**Figure 4.** Questionnaire based on the VRK + TSL learning model.

The study was reviewed and approved by the Office of the Khon Kaen University Ethics Committee in Human Research. The questionnaires were distributed to students in Grades 7–9 (13–17 years old) in two schools in the northeast of Thailand: School for the Deaf, Khon Kaen, and School for the Deaf, Udon Thani. The objectives and intended benefits of the research were explained to the 82 participants, who were also informed that there were no right or wrong answers and that their class scores would not be affected. The video was also played to aid in student understanding. The students answered the questions based solely on their preferences and were encouraged not to copy from each other.

### 3. VRK + TSL learning styles analysis using data mining (Figure 5).



**Figure 5.** The innovative procedure for classifying VRK + TSL learning styles.

This study used educational data to extract knowledge from data obtained via the questionnaire and to determine the learning styles of participants. The data were processed before being analyzed using the following steps:

- The questionnaires were collected from the deaf students at the special schools in Khon Kaen and Udon Thani, Thailand, as shown in Table 1.
- The data were then screened and any non-useful content was discarded.
- The data were classified into four learning groups based on the VRK + TSL model and converted into the format necessary for the next processing step.
- Data analysis was conducted using data mining via the decision tree, random forest, Bayesian network, naïve Bayes, multi-layer perceptron, and KNN algorithms. Feature selection was utilized to help select the right feature in the analysis, and feature selection was used with SMOTE to solve data imbalance problems.
- After analysis, models were developed and assessed for optimum suitability and effectiveness.

**Table 1.** Data set for data analysis.

Gender	Age	Level	Grade	Habitat	School before Entering	Domicile	School for the Deaf	Learning Style
Female	13	G.7	3.90	Dorm	School of the Deaf	Kalasin	Khon Kaen	V
Male	14	G.8	3.25	Dorm	School of the Deaf	Khon Kaen	Khon Kaen	K
Male	14	G.8	3.46	Dorm	School of the Deaf	Sakon Nakhon	Udon Thani	TSL
Female	16	G.9	3.00	Dorm	School of the Deaf	Udon Thani	Udon Thani	K
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

### 3.6.2. Efficacy Measurement

To test the effectiveness of each model, statistical outcomes, accuracy, precision (PREC), recall (REC), and F-measure were taken into consideration, based on the confusion matrix [56] table.

True positive (TP) occurred when the prediction was true and based on expectation.

True negative (TN) occurred when the prediction was true but not based on expectation.

False positive (FP) occurred when the prediction was false but based on expectation (i.e., the reality did not refer to a certain object, but the prediction did refer to it).

False negative (FN) occurred when the prediction was false and not based on an expectation (i.e., the reality referred to a certain object but the prediction did not).

Accuracy is calculated as the number of all correct predictions divided by the total number in the dataset. The best accuracy possible is 1.0 and the worst is 0.0, and it can be defined as

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (8)$$

PREC is calculated as the number of correct positive predictions divided by the total number of positive predictions. It is also called the positive predictive value. The best PREC is 1.0 and the worst is 0.0, and it can be defined as

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

REC is calculated as the number of correct positive predictions divided by the total number of positives. It is also called TP rate. The best sensitivity is 1.0 and the worst is 0.0, and it can be defined as

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

*F-Measure* is the process used to find the PREC and REC, using the following equation:

$$F\text{-measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

#### 4. Results

Data obtained from students in Grades 7–9 at the two Schools for the Deaf were tested for effectiveness. The results were divided into 5-fold and 10-fold validations (Table 2).

**Table 2.** Efficiency measurement of each algorithm where the data were divided into 5-folds and 10-folds.

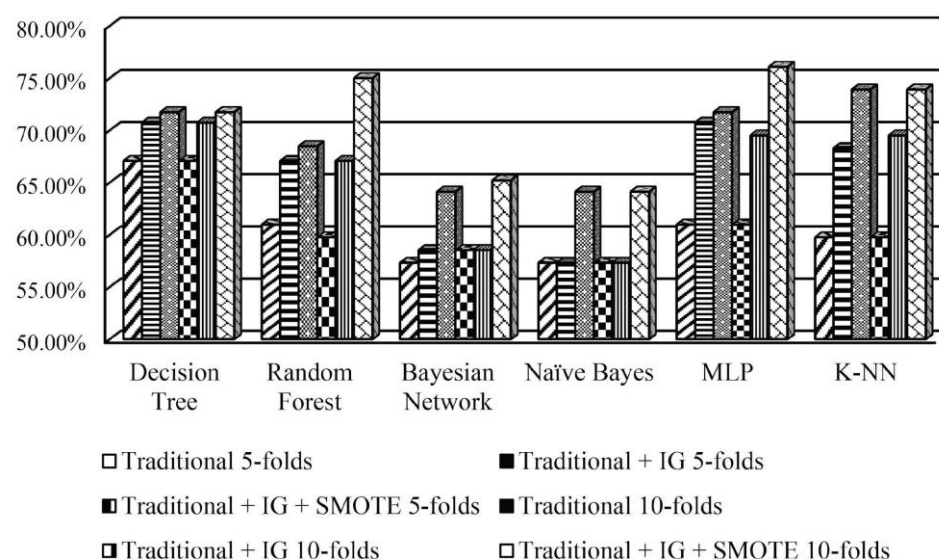
Data Division	Algorithm	TP Rate	FP Rate	Precision	Recall	F-Measure	Accuracy
5-fold	Decision Tree	0.671	0.237	0.644	0.671	0.637	67.0732%
	Decision Tree + IG	0.707	0.205	0.693	0.707	0.682	70.7317%
	Decision Tree + IG + SMOTE	0.717	0.168	0.723	0.717	0.703	71.7391%
10-fold	Decision Tree	0.671	0.237	0.644	0.671	0.637	67.0732%
	Decision Tree + IG	0.707	0.205	0.693	0.707	0.682	70.7317%
	Decision Tree + IG + SMOTE	0.717	0.168	0.723	0.717	0.703	71.7391%
5-fold	Random Forest	0.573	0.252	0.553	0.573	0.561	57.3171%
	Random Forest + IG	0.671	0.222	0.671	0.671	0.650	67.0732%
	Random Forest + IG + SMOTE	0.685	0.151	0.684	0.685	0.682	68.4783%
10-fold	Random Forest	0.598	0.249	0.546	0.598	0.569	59.7561%
	Random Forest + IG	0.671	0.223	0.630	0.671	0.639	67.0732%
	Random Forest + IG + SMOTE	0.750	0.145	0.739	0.750	0.737	75.0000%
5-fold	Bayesian Network	0.573	0.405	0.453	0.573	0.497	57.3171%
	Bayesian Network + IG	0.585	0.402	0.479	0.585	0.515	58.5366%
	Bayesian Network + IG + SMOTE	0.641	0.249	0.551	0.641	0.585	64.1304%
10-fold	Bayesian Network	0.585	0.390	0.475	0.585	0.515	58.5366%
	Bayesian Network + IG	0.585	0.427	0.492	0.585	0.512	58.5366%
	Bayesian Network + IG + SMOTE	0.652	0.264	0.567	0.652	0.590	65.2174%
5-fold	Naïve Bay	0.573	0.405	0.453	0.573	0.497	57.3171%
	Naïve Bay + IG	0.573	0.417	0.457	0.573	0.497	57.3171%
	Naïve Bay + IG + SMOTE	0.641	0.250	0.555	0.641	0.584	64.1304%
10-fold	Naïve Bay	0.573	0.417	0.467	0.573	0.501	57.3171%
	Naïve Bay + IG	0.573	0.430	0.474	0.573	0.502	57.3171%
	Naïve Bay + IG + SMOTE	0.641	0.259	0.550	0.641	0.581	64.1304%
5-fold	MLP	0.610	0.271	0.574	0.610	0.582	60.9756%
	MLP + IG	0.707	0.229	0.694	0.707	0.676	70.7317%
	MLP + IG + SMOTE	0.717	0.149	0.704	0.717	0.708	71.7391%
10-fold	MLP	0.610	0.246	0.573	0.610	0.586	60.9756%
	MLP + IG	0.695	0.231	0.649	0.695	0.659	69.5122%
	MLP + IG + SMOTE	0.761	0.141	0.752	0.761	0.745	76.0870%
5-fold	K – NN	0.598	0.311	0.553	0.598	0.562	59.7561%
	K – NN + IG	0.683	0.232	0.679	0.683	0.658	68.2927%
	K – NN + IG + SMOTE	0.739	0.149	0.745	0.739	0.722	73.9130%
10-fold	K – NN	0.598	0.311	0.553	0.598	0.562	59.7561%
	K – NN + IG	0.695	0.229	0.692	0.695	0.668	69.5122%
	K – NN + IG + SMOTE	0.739	0.149	0.745	0.739	0.722	73.9130%

Figure 6 shows the division of the data into 5-fold and 10-fold validations together with their varying accuracies for the six algorithms.

The experiments were conducted in three patterns. In the first pattern, the primary data were evaluated for accuracy using the decision tree, random forest, Bayesian network, naïve Bayes, multi-layer perceptron, and KNN classifier algorithms. In the second pattern, the primary data were processed via the feature selection method to find the factors affecting learning patterns. When the data were resynthesized, accuracy could be determined once again using the six comparative algorithms. In the third pattern, the primary data were processed via the feature selection method to find the factors affecting learning patterns.

SMOTE was used to resolve imbalanced data and resynthesize the data. Next, accuracy was tested again via the six algorithms.

The outcomes indicated that each algorithm increased accuracy by adjusting the input data and dividing the data into 5-fold and 10-fold validations. Using both these methods, the accuracy value increased for every algorithm. Dividing the data 10-fold yielded greater accuracy than 5-fold. The highest accuracy was achieved using the multi-layer perceptron algorithm and the lowest was achieved using the naïve Bayes algorithm.



**Figure 6.** Accuracy of each algorithm for 5- and 10-fold validations and additional aspects.

## 5. Discussion

This research investigated the classification used to predict the learning styles of deaf students, using the decision tree, random forest, Bayesian network, naïve Bayes, multi-layer perceptron, and KNN algorithms together with feature selection and SMOTE. Their efficiency and accuracy were evaluated using various measurement tools. The combination that achieved the highest accuracy was that of the multi-layer perceptron, random forest, and KNN algorithms, which were performed together with feature selection using information gain and SMOTE to rectify imbalanced data [83]. Classifications were conducted to predict learning styles by the selection of the factors that affect learning, and SMOTE was employed to manage and resolve imbalanced data. Oversampling randomized the minority class and balanced it with the majority class, further enhancing the effectiveness of the classification in predicting the minority class, as well as dividing the data to evaluate model effectiveness. Data types were classified into both 5- and 10-fold validations. Data were divided into  $k$  groups and where  $k$  equaled 1 to  $n$  groups, and each group contained the same amount of data. One group was then selected to be the test group, and the remaining groups acted as training groups. The test group was switched and circulated into  $k$  rounds until every rotation was complete. The techniques mentioned above were used to resolve problems and resulted in better outcomes.

## 6. Conclusions and Recommendations

### 6.1. Conclusions

The aim of this research was to classify the learning patterns of deaf students based on their learning behaviors. The data classifications were not balanced due to their differences in size. Therefore, to evaluate the effectiveness of each model, the data were divided into 5- and 10-fold validations and tested with each comparative algorithm. The algorithm with the highest accuracy was multi-layer perceptron with 60.9756% accuracy. Data accuracy was further improved via feature selection. The multi-layer perceptron algorithm with 5- and 10-fold classifications yielded accuracy rates of 70.7317% and 69.5122%, respectively.

However, imbalanced data could still exist, as the classifications were conducted with both majority and minority classes simultaneously. The data properties from the majority class could overshadow those of the minority class, which would lead to reduced effectiveness of the minority class data. To solve this problem, SMOTE was used to balance the data from both classes via a random process in conjunction with feature selection. The resynthesized process resulted in improved outcomes, namely, the 5-fold classification with multi-layer perceptron presented 71.7391% accuracy, and the 10-fold classification increased accuracy to 76.0870%.

The effectiveness of the decision tree, random forest, Bayesian network, naïve Bayes, multi-layer perceptron, and KNN algorithms was evaluated for the classification of learning patterns. The accuracy of each algorithm increased when feature selection and SMOTE were applied together. Multi-layer perceptron was the algorithm with the highest accuracy when performed in conjunction with feature selection and SMOTE. This study is the first to demonstrate that feature selection with SMOTE can improve the accuracy of an algorithm and solve imbalanced data problems.

## 6.2. Recommendations

- Further education should be based on the classification of learning styles for deaf students, focusing on areas where they have preferences, as being engaged and enjoying the educational process benefits future careers.
- Learning factors for deaf students could be further expanded, potentially leading to the discovery of other more important factors.
- The concept of the model used in this study could be applied to teaching, prediction, or instructional media for deaf students or other learners with special needs.
- The data analysis employed in this research was adapted specifically for deaf students and could be further applied to other groups with imbalanced data, for example, speech-impaired or visually impaired learners.

Finally, this research has allowed us to affirm that behavioral patterns in deaf students are very disparate; students can show different types of behavior that range from permanent involvement to virtual silence.

To maximize their potential, it would be necessary to design methodological strategies that promote active student participation and a collaborative approach to the construction of learning.

**Author Contributions:** Conceptualization, U.K.; Data curation, T.L.; Formal analysis, T.L.; Funding acquisition, U.K.; Investigation, T.L.; Methodology, U.K.; Project administration, U.K.; Resources, T.L.; Software, T.L.; Supervision, U.K.; Validation, T.L. and U.K.; Visualization, T.L.; Writing—original draft, T.L. and U.K.; Writing—review & editing, U.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Research and Academic Affairs Promotion Fund, Faculty of Science, Khon Kaen University, fiscal year 2022 (RAAPF).

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of the Khon Kaen University Ethics Committee for Human Research (Institutional Review Board Number; IRB00012791, Federal wide Assurance; FWA00003418 and date of approval: 27 January 2022).

**Informed Consent Statement:** Participation in the study was voluntary and all participants provided written, informed consent before participation.

**Data Availability Statement:** The participants' datasets generated and analyzed during the current study are available on reasonable request from the corresponding author.

**Acknowledgments:** The author would like to thank the College of Computing, Khon Kaen University for the research location and equipment used in this study. My gratitude is also extended to the teachers and students at the School for the Deaf in both Udon Thani and Khon Kaen for their valuable assistance. Further recognition goes to the National Electronics and Computer Technology Center

(NECTEC) for the production of the document that embedded the two-dimensional barcode (QR code) within the sign language videos, as well as the entire SQR (Sign Language QR Code for the Deaf) system. The authors would like to acknowledge Publication Clinic KKU, Thailand.

**Conflicts of Interest:** The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## References

- Office of the National Education Commission; Office of the Prime Minister, Thailand. *National Education Act of B.E. 2542*; Cabinet and Royal Gazette Publishing Office: Bangkok, Thailand, 1999; p. 28.
- Jiawei Han, M.K. *Data Mining: Concepts and Techniques*, 2nd ed.; Morgan Kaufmann Publishers: Burlington, MA, USA, 2006.
- Baker, R.S.J.D.; Yacef, K. The State of Educational Data Mining in 2009: A Review and Future Visions. *J. Educ. Data Min.* **2009**, *1*, 3–17.
- Romero, C.; Ventura, S.; García, E. Data mining in course management systems: Moodle case study and tutorial. *Comput. Educ.* **2008**, *51*, 368–384. [[CrossRef](#)]
- Lee, Y.J. Developing an efficient computational method that estimates the ability of students in a Web-based learning environment. *Comput. Educ.* **2012**, *58*, 579–589. [[CrossRef](#)]
- Wang, M.; Jia, H.; Sugumaran, V.; Ran, W.; Liao, J. A web-based learning system for software test professionals. *IEEE Trans. Educ.* **2011**, *54*, 263–272. [[CrossRef](#)]
- Bhattacharyya, E.; Shariff, A.B.M.S. Learning Style and its Impact in Higher Education and Human Capital Needs. *Procedia Soc. Behav. Sci.* **2014**, *123*, 485–494. [[CrossRef](#)]
- Hundhausen, C.D.; Olivares, D.M.; Carter, A.S. IDE-based learning analytics for computing education: A process model, critical review, and research agenda. *ACM Trans. Comput. Educ.* **2017**, *17*, 1–26. [[CrossRef](#)]
- Pantho, O.; Tiantong, M. Using Decision Tree C4. 5 Algorithm to Predict VARK Learning Styles. *Int. J. Comput. Internet Manag.* **2016**, *24*, 58–63.
- Paireekreng, W.; Prexawanprasut, T. An integrated model for learning style classification in university students using data mining techniques. In Proceedings of the ECTI-CON 2015–2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, Hua Hin, Thailand, 24–27 June 2015; pp. 1–5.
- Chang, Y.C.; Kao, W.Y.; Chu, C.P.; Chiu, C.H. A learning style classification mechanism for e-learning. *Comput. Educ.* **2009**, *53*, 273–285. [[CrossRef](#)]
- Petchboonmee, P.; Phonak, D.; Tiantong, M. A Comparative Data Mining Technique for David Kolb’s Experiential Learning Style Classification. *Int. J. Inf. Educ. Technol.* **2015**, *5*, 672–675.
- Ahmad, N.B.H.; Shamsuddin, S.M.; Ahmad, N.B.H.; Shamsuddin, S.M. A comparative analysis of mining techniques for automatic detection of student’s learning style. In Proceedings of the 2010 10th International Conference on Intelligent Systems Design and Applications, ISDA’10, Cairo, Egypt, 29 November–1 December 2010; pp. 877–882.
- Deo, T.Y.; Patange, A.D.; Pardeshi, S.S.; Jegadeeshwaran, R.; Khairnar, A.N.; Khade, H.S. A White-Box SVM Framework and its Swarm-Based Optimization for Supervision of Toothed Milling Cutter through Characterization of Spindle Vibrations. *arXiv* **2021**, arXiv:2112.08421.
- Patange, A.D.; Jegadeeshwaran, R. Application of bayesian family classifiers for cutting tool inserts health monitoring on CNC milling. *Int. J. Progn. Heal. Manag.* **2020**, *11*. [[CrossRef](#)]
- Patil, S.S.; Pardeshi, S.S.; Patange, A.D.; Jegadeeshwaran, R. Deep Learning Algorithms for Tool Condition Monitoring in Milling: A Review. In Proceedings of the International Virtual Conference on Intelligent Robotics, Mechatronics and Automation Systems 2021, Chennai, India, 26–27 March 2021; Volume 1969, p. 12039.
- Carter, A.S.; Hundhausen, C.D.; Adesope, O. Blending measures of programming and social behavior into predictive models of student achievement in early computing courses. *ACM Trans. Comput. Educ.* **2017**, *17*, 1–20. [[CrossRef](#)]
- Rakap, S. Impacts of learning styles and computer skills on adult students’ learning online. *Turk. Online J. Educ. Technol.* **2010**, *9*, 108–115.
- Hossain, M.M.; Abdullah, A.B.M.; Prybutok, V.R.; Talukder, M. the Impact of Learning Style on Web Shopper Electronic Catalog Feature Preference. *J. Electron. Commer. Res.* **2009**, *10*, 1–12.
- James, S.; D’Amore, A.; Thomas, T. Learning preferences of first year nursing and midwifery students: Utilising VARK. *Nurse Educ. Today* **2011**, *31*, 417–423. [[CrossRef](#)] [[PubMed](#)]
- Hung, Y.H.; Chang, R.I.; Lin, C.F. Hybrid learning style identification and developing adaptive problem-solving learning activities. *Comput. Hum. Behav.* **2016**, *55*, 552–561. [[CrossRef](#)]
- Koch, J.; Salamonson, Y.; Rolley, J.X.; Davidson, P.M. Learning preference as a predictor of academic performance in first year accelerated graduate entry nursing students: A prospective follow-up study. *Nurse Educ. Today* **2011**, *31*, 611–616. [[CrossRef](#)] [[PubMed](#)]
- Ictenbas, B.D.; Eryilmaz, H. Determining learning styles of engineering students to improve the design of a service course. *Procedia Soc. Behav. Sci.* **2011**, *28*, 342–346. [[CrossRef](#)]



24. Alkhasawneh, I.M.; Mrayyan, M.T.; Docherty, C.; Alashram, S.; Yousef, H.Y. Problem-based learning (PBL): Assessing students' learning preferences using vark. *Nurse Educ. Today* **2008**, *28*, 572–579. [CrossRef] [PubMed]
25. Jena, R.K. Predicting students' learning style using learning analytics: A case study of business management students from India. *Behav. Inf. Technol.* **2018**, *37*, 978–992. [CrossRef]
26. Felder, R.; Silverman, L. Learning and Teaching Styles in Engineering Education. *Eng. Educ.* **1988**, *78*, 674–681.
27. Kolb, D.A. Experiential Learning: Experience as The Source of Learning and Development. *Prentice Hall Inc.* **1984**, *1*, 20–38.
28. Kolb, D.A. Disciplinary inquiry norms and student learning styles: Diverse pathways for growth. *Mod. Am. Coll.* **1981**, 21–43. Available online: [https://www.researchgate.net/publication/283922529\\_Learning\\_Styles\\_and\\_Disciplinary\\_Differences](https://www.researchgate.net/publication/283922529_Learning_Styles_and_Disciplinary_Differences) (accessed on 15 March 2022).
29. Felder, R.M. Reaching the second tier. *Journal of College Science Teaching. J. Coll. Sci. Teach.* **1993**, *23*, 286–290.
30. Fleming, N.D. *Teaching and Learning Styles: VARK Strategies*; IGI Global: Christchurch, New Zealand, 2001; ISBN 0473079569.
31. Eicher, J.P. *Making the Message Clear: Communicating for Business*; Grinder Delozier & Assoc: Scotts Valley, CA, USA, 1987; ISBN 0929514009.
32. Hawk, T.F.; Shah, A.J. Using Learning Style Instruments to Enhance Student Learning. *Decis. Sci. J. Innov. Educ.* **2007**, *5*, 1–19. [CrossRef]
33. Fleming, N.D. I'm different; not dumb. Modes of presentation (VARK) in the tertiary classroom. In Proceedings of the Research and Development in Higher Education, Proceedings of the Annual Conference of the Higher Education and Research Development Society of Australasia, Rockhampton, QLD, Australia, 4–7 July 1995; 1995; Volume 18, pp. 308–313.
34. Fleming, N.; Baume, D. *Learning Styles Again: Varking Up the Right Tree! Educational Developments*; SEDA Ltd: Blackwood, UK, 2006; Volume 7, pp. 4–7.
35. Al-Jumeily, D.; Hussain, A.; Alghamdi, M.; Dobbins, C.; Lunn, J. Educational crowdsourcing to support the learning of computer programming. *Res. Pract. Technol. Enhanc. Learn.* **2015**, *10*, 13. [CrossRef]
36. Alkhasawneh, E. Using VARK to assess changes in learning preferences of nursing students at a public university in Jordan: Implications for teaching. *Nurse Educ. Today* **2013**, *33*, 1546–1549. [CrossRef]
37. Liew, S.C.; Sidhu, J.; Barua, A. The relationship between learning preferences (styles and approaches) and learning outcomes among pre-clinical undergraduate medical students Approaches to teaching and learning. *BMC Med. Educ.* **2015**, *15*, 44. [CrossRef]
38. Drago, W.A.; Wagner, R.J. Vark preferred learning styles and online education. *Manag. Res. News* **2004**, *27*, 1–13. [CrossRef]
39. Dutsinma, L.I.F.; Temdee, P. VARK Learning Style Classification Using Decision Tree with Physiological Signals. *Wirel. Pers. Commun.* **2020**, *115*, 2875–2896. [CrossRef]
40. Truong, H.M. Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities. *Comput. Hum. Behav.* **2016**, *55*, 1185–1193. [CrossRef]
41. Jones, M.; Oviatt, S.; Brewster, S.; Penn, G.; Munteanu, C.; Whittaker, S.; Rajput, N.; Nanavati, A. We Need to Talk: HCI and the Delicate Topic of Spoken Language Interaction. In Proceedings of the Conference on Human Factors in Computing Systems—Proceedings, Paris, France, 27 April–2 May 2013; Volume 2013, pp. 2459–2464.
42. Matthews, T.; Carter, S.; Pai, C.; Fong, J.; Mankoff, J. Scribe4Me: Evaluating a Mobile Sound Transcription Tool for the Deaf. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Proceedings of the International Conference on Ubiquitous Computing, Orange County, CA, USA, 17–21 September 2006*; Springer: Amsterdam, The Netherlands, 2006; Volume 4206 LNCS, pp. 159–176.
43. Kafle, S.; Huenerfauth, M. Predicting the understandability of imperfect english captions for people who are deaf or hard of hearing. *ACM Trans. Access. Comput.* **2019**, *12*, 1–32. [CrossRef]
44. Nanayakkara, S.C.; Wyse, L.; Ong, S.H.; Taylor, E.A. Enhancing musical experience for the hearing-impaired using visual and haptic displays. *Hum. Comput. Interact.* **2013**, *28*, 115–160.
45. Jacobs, L.M. *A Deaf Adult Speaks Out*; Gallaudet University Press: Chicago, IL, USA, 1974.
46. Stokoe, W.C.; Marschark, M. Sign language structure: An outline of the visual communication systems of the american deaf. *J. Deaf. Stud. Deaf. Educ.* **2005**, *10*, 3–37. [CrossRef]
47. Liddell, S.K. *American Sign Language*; Mouton De Gruyter: Berlin, Germany, 1980.
48. De Guilhermino Trindade, D.F.; Guimares, C.; Antunes, D.R.; Sánchez Garcia, L.; da Lopes, S.R.A.; Fernandes, S. Challenges of knowledge management and creation in communities of practice organisations of Deaf and non-Deaf members: Requirements for a Web platform. *Behav. Inf. Technol.* **2012**, *31*, 799–810. [CrossRef]
49. LeMaster, B.; Monaghan, L. Variation in Sign languages. In *A companion to Linguist*; Alessandro Duranti, Blackwell: Carlton, Australia, 2004; pp. 141–165.
50. Zeshan, U. Sign Languages of the World. In *Encyclopedia of Language & Linguistics*; Brown, K., Ed.; Elsevier: Oxford, UK, 2006; pp. 358–365.
51. Nonaka, A.M. (Almost) everyone here spoke Ban Khor Sign Language—Until they started using TSL: Language shift and endangerment of a Thai village sign language. *Lang. Commun.* **2014**, *38*, 54–72. [CrossRef]
52. Zhao, Y.; Sun, P.; Xie, R.; Chen, H.; Feng, J.; Wu, X. The relative contributions of phonological awareness and vocabulary knowledge to deaf and hearing children's reading fluency in Chinese. *Res. Dev. Disabil.* **2019**, *92*, 103444. [CrossRef]
53. Webster, A. *Deafness, Development and Literacy*; Methuen & Co. Ltd: London, UK, 2017; ISBN 9781351236010.

54. Myklebust, H.R. *The Psychology of Deafness: Sensory Deprivation, Learning, and Adjustment*; Grune and Stratton: New York, NY, USA, 1964; Volume 2.
55. Berry, M.J.A.; Linoff, G.S. *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*; John Wiley & Sons: Hoboken, NJ, USA, 2004; ISBN 0471179809.
56. Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques: Concepts and Techniques*, 3rd ed.; Elsevier: Amsterdam, The Netherlands, 2012; ISBN 978-0-12-381479-1.
57. Sharma, T.C.; Jain, M. WEKA Approach for Comparative Study of Classification Algorithm. *Int. J. Adv. Res. Comput. Commun. Eng.* **2013**, *2*, 1925–1931.
58. Lin, C.F.; Yeh, Y.C.; Hung, Y.H.; Chang, R.I. Data mining for providing a personalized learning path in creativity: An application of decision trees. *Comput. Educ.* **2013**, *68*, 199–210. [[CrossRef](#)]
59. Wu, X.; Kumar, V.; Ross, Q.J.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G.J.; Ng, A.; Liu, B.; Yu, P.S.; et al. Top 10 algorithms in data mining. *Knowl. Inf. Syst.* **2008**, *14*, 1–37. [[CrossRef](#)]
60. Amatriain, X.; Pujol, J.M. Data mining methods for recommender systems. In *Recommender Systems Handbook*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2015; pp. 227–262. ISBN 9781489976376.
61. Yin, H.H.S.; Langenheldt, K.; Harlev, M.; Mukkamala, R.R.; Vatrappu, R. Regulating Cryptocurrencies: A Supervised Machine Learning Approach to De-Anonymizing the Bitcoin Blockchain. *J. Manag. Inf. Syst.* **2019**, *36*, 37–73. [[CrossRef](#)]
62. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
63. Wang, Z.; Jiang, C.; Zhao, H.; Ding, Y. Mining semantic soft factors for credit risk evaluation in Peer-to-Peer lending. *J. Manag. Inf. Syst.* **2020**, *37*, 282–308. [[CrossRef](#)]
64. Grushka-Cockayne, Y.; Jose, V.R.R.; Lichtendahl, K.C. Ensembles of overfit and overconfident forecasts. *Manag. Sci.* **2017**, *63*, 1110–1130. [[CrossRef](#)]
65. Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* **2002**, *2*, 18–22.
66. Friedman, N.; Geiger, D.; Goldszmidt, M. Bayesian Network Classifiers. *Mach. Learn.* **1997**, *29*, 131–163. [[CrossRef](#)]
67. Ko, Y. How to use negative class information for Naive Bayes classification. *Inf. Process. Manag.* **2017**, *53*, 1255–1268. [[CrossRef](#)]
68. Lee, C.H. An information-theoretic filter approach for value weighted classification learning in naive Bayes. *Data Knowl. Eng.* **2018**, *113*, 116–128. [[CrossRef](#)]
69. Cover, T.M.; Hart, P.E. Nearest Neighbor Pattern Classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [[CrossRef](#)]
70. Zhang, M.-L.; Zhou, Z.-H. A k-nearest neighbor based algorithm for multi-label classification. In Proceedings of the 2005 IEEE International Conference on Granular Computing IEEE, Beijing, China, 25–27 July 2005; Volume 2, pp. 718–721.
71. McCulloch, W.S.; Pitts, W.H. A logical calculus of the ideas immanent in nervous activity. In *Systems Research for Behavioral Science: A Sourcebook*; Springer: Berlin/Heidelberg, Germany, 2017; Volume 5, pp. 93–96. ISBN 9781351487214.
72. Mosier, C.I.I. Problems and Designs of Cross-Validation 1. *Educ. Psychol. Meas.* **1951**, *11*, 5–11. [[CrossRef](#)]
73. Stone, M. Cross-Validatory Choice and Assessment of Statistical Predictions (With Discussion). *J. R. Stat. Soc. Ser. B* **1974**, *36*, 111–133. [[CrossRef](#)]
74. Tibshirani, R.J.; Tibshirani, R. A bias correction for the minimum error rate in cross-validation. *Ann. Appl. Stat.* **2009**, *3*, 822–829. [[CrossRef](#)]
75. Chen, K.; Lei, J. Network Cross-Validation for Determining the Number of Communities in Network Data. *J. Am. Stat. Assoc.* **2018**, *113*, 241–251. [[CrossRef](#)]
76. Liu, Y.; Liao, S. Granularity selection for cross-validation of SVM. *Inf. Sci.* **2017**, *378*, 475–483. [[CrossRef](#)]
77. Yan, L.; He, Y.; Qin, L.; Wu, C.; Zhu, D.; Ran, B. A novel feature extraction model for traffic injury severity and its application to Fatality Analysis Reporting System data analysis. *Sci. Prog.* **2020**, *103*, 0036850419886471. [[CrossRef](#)]
78. Pal, N.R.; Pal, S.K. Entropy: A New Definition and its Applications. *IEEE Trans. Syst. Man Cybern.* **1991**, *21*, 1260–1270. [[CrossRef](#)]
79. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [[CrossRef](#)]
80. Rovinelli, R.J.; Hambleton, R.K. On the Use of Content Specialists in the Assessment of Criterion-Referenced Test Item Validity. *Dutch J. Educ. Res.* **1977**, *2*, 49–60.
81. Krishnamoorthy, D.; Lokesh, D. Process of building a dataset and classification of vark learning styles with machine learning and predictive analytics models. *J. Contemp. Issues Bus. Gov.* **2021**, *26*, 903–910.
82. Likert, R. A technique for the measurement of attitudes. *Arch. Psychol.* **1932**, *22*, 55.
83. Ali, H.; Salleh, M.N.M.; Saedudin, R.; Hussain, K.; Mushtaq, M.F. Imbalance class problems in data mining: A review. *Indones. J. Electr. Eng. Comput. Sci.* **2019**, *14*, 1560–1571. [[CrossRef](#)]