



## Article

# Intelligent Transmit Antenna Selection Schemes for High-Rate Fully Generalized Spatial Modulation

Hindavi Kishor Jadhav <sup>1</sup>, Vinoth Babu Kumaravelu <sup>1,\*</sup>, Arthi Murugadass <sup>2,\*</sup>,  
Agbotiname Lucky Imoize <sup>3,4</sup>, Poongundran Selvaprabhu <sup>1</sup> and Arunkumar Chandrasekhar <sup>5</sup>

- <sup>1</sup> Department of Communication Engineering, School of Electronics Engineering, Vellore Institute of Technology, Vellore 632014, India; hindavikj@gmail.com (H.K.J.); poongundrano7@gmail.com (P.S.)
- <sup>2</sup> Department of Computer Science and Engineering (AI & ML), Sreenivasa Institute of Technology and Management Studies, Chittoor 517127, India
- <sup>3</sup> Department of Electrical and Electronics Engineering, Faculty of Engineering, University of Lagos, Lagos 100213, Nigeria; aimoize@unilag.edu.ng
- <sup>4</sup> Department of Electrical Engineering and Information Technology, Institute of Digital Communication, Ruhr University, 44801 Bochum, Germany
- <sup>5</sup> Department of Sensor and Biomedical Technology, School of Electronics Engineering, Vellore Institute of Technology, Vellore 632014, India; arunkumar.c@vit.ac.in
- \* Correspondence: vinothbab@gmail.com (V.B.K.); arthimdas@gmail.com (A.M.)

**Abstract:** The sixth-generation (6G) network is supposed to transmit significantly more data at much quicker rates than existing networks while meeting severe energy efficiency (EE) targets. The high-rate spatial modulation (SM) methods can be used to deal with these design metrics. SM uses transmit antenna selection (TAS) practices to improve the EE of the network. Although it is computationally intensive, free distance optimized TAS (FD-TAS) is the best for performing the average bit error rate (ABER). The present investigation aims to examine the effectiveness of various machine learning (ML)-assisted TAS practices, such as support vector machine (SVM), naïve Bayes (NB),  $K$ -nearest neighbor (KNN), and decision tree (DT), to the small-scale multiple-input multiple-output (MIMO)-based fully generalized spatial modulation (FGSM) system. To the best of our knowledge, there is no ML-based antenna selection schemes for high-rate FGSM. SVM-based TAS schemes achieve ~71.1% classification accuracy, outperforming all other approaches. The ABER performance of each scheme is evaluated using a higher constellation order, along with various transmit antennas to achieve the target ABER of  $10^{-5}$ . By employing SVM for TAS, FGSM can achieve a minimal gain of ~2.2 dB over FGSM without TAS (FGSM-NTAS). All TAS strategies based on ML perform better than FGSM-NTAS.

**Keywords:** free distance optimized transmit antenna selection (FD-TAS); fully generalized spatial modulation (FGSM); machine learning (ML); support vector machine (SVM); transmit antenna selection (TAS)



**Citation:** Jadhav, H.K.; Kumaravelu, V.B.; Murugadass, A.; Imoize, A. L.; Selvaprabhu, P.; Chandrasekhar, A. Intelligent Transmit Antenna Selection Schemes for High-Rate Fully Generalized Spatial Modulation. *Future Internet* **2023**, *15*, 281. <https://doi.org/10.3390/fi15080281>

Academic Editor: Claude Chaudet

Received: 2 July 2023

Revised: 7 August 2023

Accepted: 17 August 2023

Published: 21 August 2023



**Copyright:** © 2023 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

In comparison to fifth-generation (5G) wireless communication networks, 6G networks are expected to have much higher spectral, energy, and cost efficiency, with higher data rates (in Tbps), latency reduced by a factor of ten, connection density increased by a factor of a hundred, and increased intelligence for full automation. Innovations such as index modulation (IM), reconfigurable intelligent surfaces (RISs), non-orthogonal multiple access (NOMA), and artificial intelligence (AI) will all be introduced into 6G networks to fulfill these overall objectives [1–6]. Massive MIMO designs can increase the spectral efficiency (SE) of 6G systems because of the higher frequencies and dense networks. In 6G, technologies like SM, RIS, NOMA, orbital angular momentum, and rate splitting multiple access have the potential to increase SE [7].

The energy efficiency (EE) of wireless networks may be enhanced by using four major approaches: (1) Resource allocation. Rather than maximizing throughput, the EE of the network is optimized by efficiently distributing resources [8]. (2) Instead of simply increasing the coverage area, the second option employs infrastructure nodes to increase the network's EE. (3) Use renewable energy sources to power communication systems. (4) Design energy-efficient hardware solutions. In a network, IM and its derivatives can also be utilized to shoot up EE [9,10].

Massive MIMO is an essential technique for achieving the SE requirements of 5G. When adopting a multi-antenna system, two issues arise: inter-antenna synchronization (IAS) and inter-channel interference (ICI) [11]. With more radio frequency (RF) links, the hardware complexity of the system grows, lowering EE [12]. To overcome this problem, the concept of SM has been presented [9,10]. Because one antenna is functional at a time, the RF chain is simpler to handle in SM [9–11].

SM's SE is given by

$$\alpha_{SM} = \log_2(M) + \log_2(N_{tx}), \quad (1)$$

where the number of transmit antennas and modulation order are denoted by  $N_{tx}$  and  $M$ , respectively. The SE of SM is proportional to  $\log_2(N_{tx})$ . As a result, in order to enhance SE, a higher proportion of transmit antennas is required. SM cannot satisfy SE demands since only one antenna is simultaneously active [11]. High-rate SM variants are needed for future-generation networks to meet SE requirements [13]. The high-rate SM variants may transmit the same or distinct symbols at the same time across multiple antennas, depending on their operating principle. FGSM is a cutting-edge SM variant in which the transmitter uses just one or several, or all of its antennas for transmission [13]. As a result, the possible SE grows linearly with  $N_{tx}$ . Gudla V.V. et al. briefly discussed the system architecture and functioning theory of FGSM [13].

Transmit diversity gain could not be achieved in SM since only one antenna is fully functional at once [14–17]. The integration of TAS practices into SM improves its transmit diversity. The receiver decides on transmitter antenna subsets driven by the channel quality information (CQI). In ML, machines are taught new skills so that they may execute tasks on their own using data. In terms of planning and optimizing future-generation networks, ML plays a critical role. ML may be used to solve a wide range of technical difficulties in future-generation systems, including massive MIMO, NOMA, device-to-device (D2D) networks, etc. [2,3]. The present research performs TAS on FGSM using ML algorithms, and the effectiveness of these algorithms is evaluated using classification accuracy and ABER.

#### *Related Work*

The future-generation network demands improvements in EE and SE [1]. The SE of SM is proportional to  $\log_2(N_{tx})$ , which lowers the performance of SM in terms of SE [9,10]. To overcome this, high-rate versions of SM are developed [13]. FGSM is a high-rate variation of SM, where transmit antennas and SE are linearly related and could meet the expanding demands of 6G [13].

Lower-powered integrated circuits, antenna diversity, or a mix of the two, may be adopted to boost EE [18]. The traditional SM technique does not provide transmit diversity gain. TAS practices have been incorporated to boost the SM system's transmit diversity gain [14]. TAS approaches are applied to overcome the cost and complexity challenges in massive MIMO systems [12]. Different massive MIMO configurations have shown that antenna selection can improve performance and lower RF costs [19]. For optimal channel utilization, a novel SM technique is developed by fusing transmit mode switching and adaptive modulation techniques [20]. The free distance (FD) strategy is computationally more expensive due to the requirement of simultaneously searching for antenna pairs and constellation orders that maximize the least FD. Although FD-TAS provides the best EE performance, its implementation is more challenging. The FD-TAS system has gained popularity and is used as a benchmark in many TAS-related articles [3,13–17,21].

A capacity-optimized antenna selection (COAS) technique is addressed in conjunction with FD-TAS, where the antenna subset with the most prominent channel capacity is chosen [13–15,17]. The correlation angle of two antennas is used to evaluate the TAS [13,16]. In this study, the antenna subset with the least correlation is prioritized. The capacity and correlation angle-based technique is proposed to enhance ABER performance [13] while increasing computing complexity. A partitioning strategy that relies on the capacity and correlation angle further reduces the system's complexity [13]. Sub-optimal techniques are less efficient than FD-TAS when it comes to TAS in SM and its variants [13–17].

To save the computations offered by FD-TAS, nowadays data-driven approaches are utilized. TAS is a classification problem that can be solved using supervised ML algorithms. A pattern recognition-based approach is proposed for traditional MIMO [22]. This work is tested for 5000 samples with KNN and SVM algorithms. This work can be extended further to various supervised learning algorithms and larger datasets to boost the overall system performance. Two unique features of the channel space are the element norm of  $\mathbf{H}$  and element norm of  $\mathbf{H}^H\mathbf{H}$ , which are used to analyze the performance of these algorithms. The absolute of elements of  $\mathbf{H}$  is used to analyze NB- and SVM-based TAS algorithms in conventional MIMO architecture [23]. The purpose of this research is to demonstrate how ML can be utilized to enhance the security of MIMO architecture. In the future, other ML algorithms could be employed to carry out this research. In this example, the algorithms are only trained on 10,000 samples; nevertheless, this number can be increased to improve overall efficiency.

To solve power assignment problems in adaptive SM-MIMO, supervised ML algorithms and deep neural networks (DNNs) are proposed and implemented [24]. The ABER performance in this work can be enhanced by extending the training data beyond 2000 feature label pairs. Other features, such as the angle of elements of  $\mathbf{H}$ , real and imaginary parts of elements of  $\mathbf{H}$ , etc., can be used as attributes to obtain better results. TAS as a classification problem is addressed using DNN and SVM for SM [25]. To construct the models, the channel gain and correlation properties of the column space of  $\mathbf{H}$  are retrieved. In this work, the comparison with other ML-based algorithms is not conducted with DNN and SVM. Altın, G. and Arslan, İ.A. selected both the transmit and receive antennas for SM using deep learning (DL) architectures [26]. We used DL-based algorithms to execute TAS for FGSM in [3] to meet the SE requirements and improve classification accuracy. The following is a list of research gaps identified:

- Most of these data-driven approaches are proposed and implemented for conventional MIMO and basic SM architectures.
- These data-driven strategies are neither proposed nor implemented for high-rate variants of SM.
- The TAS dataset is not available for FGSM in any of the repositories.

Listed below are the key contributions of this work:

- Through the repeated application of the FD-TAS algorithm, 4 different datasets (with a total of 10,000 entries) are produced. The dataset contains channel information for a variety of MIMO setups and antenna counts, as well as essential performance measures, like FD.
- For FGSM, four distinct supervised ML classification methods are used and tested for TAS: SVM, NB, KNN, and DT. The dataset size, the value of  $k$ , which is used in  $k$ -fold cross-validation, and parameters, such as the kernel type and hyperparameter  $K$ , are optimized to obtain the best results.
- The ABER and computational complexity of ML-based TAS approaches are contrasted to those of FGSM-NTAS and FD-TAS.
- An analysis of the proposed work is carried out in terms of SNR gain and classification accuracy.

The following is how the manuscript is structured: The principle of operation of FGSM is discussed elaborately in Section 2. The various ML algorithms used for TAS in the FGSM system are covered in Section 3. The proposed ML-based TAS practices and the conventional FD-TAS practice are subjected to a computational complexity study in Section 4. The ABER performance of ML-based algorithms is contrasted to the FD-TAS scheme in Section 5. Section 6 concludes the work by laying out future research directions.

### 2. System Model of FGSM Transceiver

FGSM is a high-rate version of SM that allows one, many, or all antennas to transmit the same symbol at the same instant. The SE is directly proportional to  $N_{tx}$  as compared to SM, which improves the SE. This variation does not require that the number of transmit antennas be a power of two. The SE of FGSM is calculated as follows:

$$\alpha_{FGSM} = \log_2(M) + (N_{tx} - 1). \tag{2}$$

Consider the input bits [0 0 1 1] to comprehend FGSM mapping. An example with an SE of 4 bpcu ( $M = 4$  and  $N_{tx} = 3$ ) is used to demonstrate FGSM mapping. As indicated in Table 1, the symbol  $s_1$  is identified using the first  $\log_2(M)$  bits, i.e., [0 0]. It will be transmitted by the antenna pair (1, 2), which is selected using the next  $(N_{tx} - 1)$  bits, i.e., [1 1], as illustrated in Table 2. For the given block of bits, the transmit vector generated is given by  $\chi = [-\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}j, -\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}j, 0]^T$ . Consider an alternative set of bits [0 1 0 1], where the symbol  $s_2$  is chosen by using the first  $\log_2(M)$  bits [0 1], as shown in Table 1. It will be transmitted by antenna 2, which is selected using the next  $(N_{tx} - 1)$  bits, i.e., [0 1], as illustrated in Table 2. For this set of bits, the vector to be transmitted is  $\chi = [0, -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}j, 0]^T$ .

**Table 1.** A sample mapping of the constellation bits for  $M = 4$ .

Data Bits	Index of Symbols	Transmitted Symbols
00	$s_1$	$-\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}j$
01	$s_2$	$-\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}j$
10	$s_3$	$\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}j$
11	$s_4$	$\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}j$

**Table 2.** A sample mapping of the spatial bits for  $N_{tx} = 3$ .

Antenna Bits	Antenna Index
00	1
01	2
10	3
11	1,2

The multipath channel  $H$  and AWGN  $\eta$  influence the transmit vector  $\chi$ . At the receiver end, the signal vector  $y$  is stated as follows:

$$\underbrace{y}_{N_{rx} \times 1} = \underbrace{H}_{N_{rx} \times N_{tx}} \underbrace{\chi}_{N_{tx} \times 1} + \underbrace{\eta}_{N_{rx} \times 1}. \tag{3}$$

Consider the MIMO configuration ( $N_{tx} = 3, N_{rx} = 4$ ) with  $M = 4$ . For the single antenna operational configuration explained earlier, the received signal vector could be expressed as follows:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \\ h_{41} & h_{42} & h_{43} \end{bmatrix} \begin{bmatrix} 0 \\ s_2 \\ 0 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}. \tag{4}$$

It comes down to

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} h_{12} \\ h_{22} \\ h_{32} \\ h_{42} \end{bmatrix} s_2 + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}. \tag{5}$$

For single antenna operational case, the generalized form of the received signal is

$$\mathbf{y} = \mathbf{h}_j s_m + \boldsymbol{\eta}, j \in \{1, 2, \dots, N_{tx}\}, m \in \{1, 2, \dots, M\}. \tag{6}$$

For the double antenna operational configuration explained earlier, the received signal vector could be indicated as follows:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \\ h_{41} & h_{42} & h_{43} \end{bmatrix} \begin{bmatrix} s_1 \\ s_1 \\ 0 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}. \tag{7}$$

It reduces to

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} h_{11} \\ h_{21} \\ h_{31} \\ h_{41} \end{bmatrix} s_1 + \begin{bmatrix} h_{12} \\ h_{22} \\ h_{32} \\ h_{42} \end{bmatrix} s_1 + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}. \tag{8}$$

For the double antenna operational case, the generalized form of the received signal is given by

$$\mathbf{y} = (\mathbf{h}_{j_1} + \mathbf{h}_{j_2}) s_m + \boldsymbol{\eta}, j_1 \neq j_2, j_1, j_2 \in \{1, 2, \dots, N_{tx}\}, m \in \{1, 2, \dots, M\}. \tag{9}$$

For the  $N_A$  number of functional antennas, the generalized form of the received signal is represented as follows:

$$\mathbf{y} = \left( \sum_{a=1}^{N_A} \mathbf{h}_{j_a} \right) s_m + \boldsymbol{\eta}, j_a \in \{1, 2, \dots, N_{tx}\}, m \in \{1, 2, \dots, M\}. \tag{10}$$

When a receiver possesses perfect CQI, the maximum likelihood detection for the FGSM system with  $N_A$  functional antennas is given by

$$[\hat{j}_{a,a=1,2,\dots,N_A}, \hat{s}_m] = \arg \min_{j_a, a=1,2,\dots,N_A, s_m} \left\| \mathbf{y} - \left( \sum_{a=1}^{N_A} \mathbf{h}_{j_a} \right) s_m \right\|^2. \tag{11}$$

### 3. Antenna Selection Schemes for FGSM

TAS for FGSM is schematically depicted in Figure 1. TAS drives up the transmit diversity in traditional MIMO and SM, as described in the literature review section. Let  $\mathbf{H}_T = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_{N_T}] \sim \mathcal{CN}(0, I_{N_{rx} \times N_T})$  be the channel matrix for a MIMO. Here,  $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_{N_T}$  are the column vectors of  $\mathbf{H}_T$ . The TAS scheme is employed to select a subset from the candidate set  $S = \{S_1, S_2, \dots, S_l, \dots, S_{N_s}\}$ .  $N_{tx}$  antennas are picked from  $N_T$  antennas when

the CQI is perfectly known to the user equipment (UE). There are  $N_s$  possible antenna subsets, each with  $N_{tx}$  antennas, as given below:

$$\begin{aligned}
 S_1 &= \{1, 2, \dots, N_{tx}\}, \\
 S_2 &= \{1, 2, \dots, N_{tx} - 1, N_{tx} + 1\}, \\
 &\vdots \\
 S_{N_s} &= \{N_T - N_{tx} + 1, \dots, N_T\},
 \end{aligned}
 \tag{12}$$

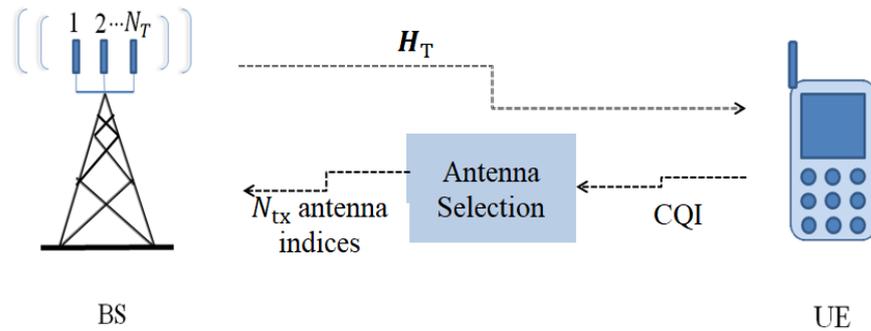
where  $S_l$  represents the  $l$ th possible transmit antenna subset. A low-rate feedback channel communicates the chosen subset antenna index to the base station (BS). The CQI feedback latency is assumed to be insignificant in this case. If  $S_l$  is selected, then the signal received is expressed as follows:

$$\mathbf{y} = \mathbf{H}_T^l \boldsymbol{\chi} + \boldsymbol{\eta},
 \tag{13}$$

where  $\mathbf{H}_T^l$  is the channel matrix corresponding to the selected set  $S_l$ . An example of the TAS candidate set mapping is illustrated in Table 3 for  $N_T = 4, N_{tx} = 3, \&N_{rx} = 4$ .

**Table 3.** The TAS candidate set mapping for  $N_T = 4, N_{tx} = 3, N_{rx} = 4$ .

TAS Candidate Set	Possible Antenna Subset
$S_1$	1,2,3
$S_2$	1,2,4
$S_3$	1,3,4
$S_4$	2,3,4



**Figure 1.** Pictorial representation of TAS for FGSM.

### 3.1. FD-TAS Based FGSM System

The purpose of FD-TAS is to find antenna pairs that optimize the least FD between the received constellations [3,13–17,21]. FD-TAS results in  $\mathbf{H}$  of size  $N_{rx} \times N_{tx}$  from  $\mathbf{H}_T$  of size  $N_{rx} \times N_T$ . FD-TAS involves the following major steps:

Step 1: Compute the number of possible subsets  $N_s = \binom{N_T}{N_{tx}}$ .

Step 2: Determine all feasible transmit vectors ( $2^{N_{FGSM}}$ ) for every antenna subset.

Step 3: Determine the minimum FD for every antenna subset using

$$d_{min}(\mathbf{H}_T) = \min_{\boldsymbol{\chi}_i \neq \boldsymbol{\chi}_j} \|\mathbf{H}_T(\boldsymbol{\chi}_i - \boldsymbol{\chi}_j)\|^2,
 \tag{14}$$

here,  $\boldsymbol{\chi}_i, \boldsymbol{\chi}_j \in \mathbb{X}$  is the set of all feasible transmit vectors.

Step 4: Find an antenna subset that has the maximum FD.

$$\hat{l} = \arg \max_{l \in \{1, 2, \dots, N_s\}} d_{min}(\mathbf{H}_T^l), \tag{15}$$

here,  $\mathbf{H}_T^l = [\mathbf{H}_T^1, \mathbf{H}_T^2, \dots, \mathbf{H}_T^{N_s}]$ .

Based on the antenna subset index obtained in step 4, the antenna indices and the associated channel matrix of size  $N_{rx} \times N_{tx}$  are computed.

### 3.2. TAS Based on ML for FGSM

In this work, TAS in FGSM is carried out with the use of ML-based methods. An ML-based classification architecture is depicted in Figure 2. At first, source data are normalized using the mean normalization technique. Pre-processed data are then used to extract features. Several ML methods are then used to train the model using the chosen features and class labels. A trained model can be used to predict the class labels for new data using data that have not previously been fed into the system. The process for the execution of ML-based TAS for FGSM is shown in Figure 3. If CQI is present at the UE, the class labels obtained by the ML algorithms are utilized to choose  $N_{tx}$  antennas from  $N_T$ . Algorithm 1 shows the processes that must be followed for ML-based TAS. This section also includes details on the dataset development process and the supervised algorithms that are employed.

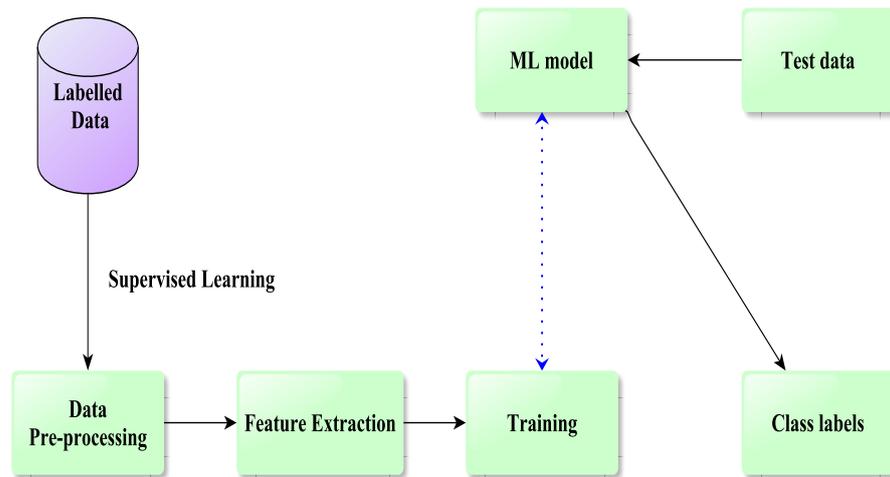


Figure 2. Workflow of supervised learning algorithms during the training and testing process.

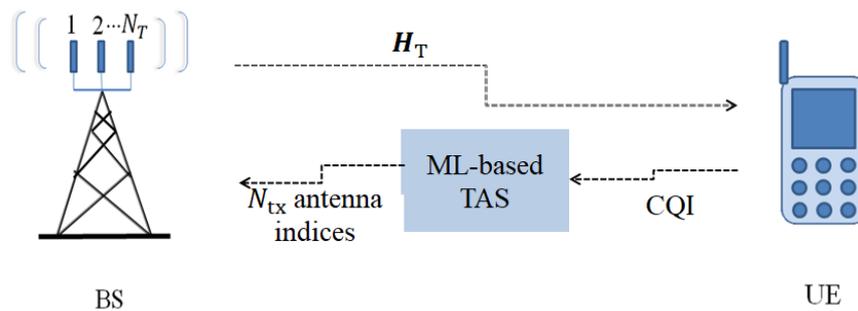


Figure 3. ML-supported TAS for FGSM.

---

**Algorithm 1** Primary steps involved in ML-based TAS

---

1. Establish a mapping between the antenna subsets and class labels.
  2. Pre-process the input data using the mean normalization technique.
  3. Identify the key performance indicators (KPI) for each input sample. Find out the class label  $c$  on the basis of the KPI.
  4. Employ the normalized feature matrix  $\hat{\mathbf{H}}_T$  and the class label  $c$  to create a learning system utilizing SVM, NB, KNN, and DT algorithms.
  5. The learning system converts the test channel  $\mathbf{H}_T^{new}$  to a normalized feature vector ( $f^{new}$ ) and determines the class label  $c$ .
  6. Compare the class label acquired in Step 5 with the map created in Step 1 to identify the antenna subset.
- 

3.3. Generation of a Dataset

Elements of channel matrices are used to generate the training dataset. A multiclass classification technique is then used to divide the channel matrix into numerous corresponding classes, each of which represents the optimal antenna subset. ML methods are used to build the model. The developed models can also be used to predict the class labels of test data.

Four major steps are involved in the collection of the dataset: (1) Generation of training data. (2) Extraction of feature vectors. (3) Evaluation of KPI. (4) Use of the KPI information to obtain the class labels.

1. Generation of training data.

To form the training data, the  $N$  number of channel matrices of size  $N_{rx} \times N_T$  are randomly generated.

$$\mathbb{H} = [\mathbf{H}_T^1, \mathbf{H}_T^2, \dots, \mathbf{H}_T^N]. \tag{16}$$

2. Extraction of feature vectors

The training performance is generally influenced by the feature vectors. In this work, four different features, i.e.,  $|h_{i,j}^n|$ ,  $\angle(h_{i,j}^n)$ , channel gain  $|h_{i,j}^n|^2$ , and channel gain of every element of the correlation matrix  $|(\mathbf{H}_T^n)^H(\mathbf{H}_T^n)|^2$ , are considered to train the models. From each sample  $\mathbf{H}_T$ , a feature of size  $N_Q$  ( $N_{rx} \times N_T$ ) is extracted. The process of extracting these four features is as follows:

- (a)  $|h_{i,j}^n|$ : The absolute value of each element of the channel matrix is considered as a dominating feature, which is used throughout this work. The extraction of the feature vector, in this case, is given as follows:

$$f^n = [|h_{1,1}^n|, \dots, |h_{1,N_T}^n|, \dots, |h_{i,j}^n|, \dots, |h_{N_{rx},N_T}^n|], n \in \{1, 2, \dots, N\}, \tag{17}$$

where  $h_{i,j}^n$  is a  $(i, j)$ th complex element of  $\mathbf{H}_T^n$ , whose absolute value is calculated as follows:

$$|h_{i,j}^n| = \sqrt{\Re(h_{i,j}^n)^2 + \Im(h_{i,j}^n)^2}. \tag{18}$$

- (b)  $\angle(h_{i,j}^n)$ : This is the second feature that is extracted to build the model from the elements of  $\mathbf{H}_T$ . The feature vector is constructed as follows:

$$f^n = [\angle(h_{1,1}^n), \dots, \angle(h_{1,N_T}^n), \dots, \angle(h_{i,j}^n), \dots, \angle(h_{N_{rx},N_T}^n)], n \in \{1, 2, \dots, N\}, \tag{19}$$

where

$$\angle(h_{i,j}^n) = \tan^{-1} \left( \frac{\Im(h_{i,j}^n)}{\Re(h_{i,j}^n)} \right). \tag{20}$$

- (c) Channel gain  $|h_{i,j}^n|^2$ : The channel gain of an individual element of the matrix is considered to be another feature to train the models whose feature vector is given by:

$$f^n = \left[ |h_{1,1}^n|^2, \dots, |h_{1,N_T}^n|^2, \dots, |h_{i,j}^n|^2, \dots, |h_{N_{rx},N_T}^n|^2 \right], n \in \{1, 2, \dots, N\}. \quad (21)$$

- (d) Squared element norm of  $(\mathbf{H}_T^n)^H (\mathbf{H}_T^n)$ : In addition, the similarity between two distinct column vectors is a major feature. Similar column vectors introduce detection ambiguity while estimating the transmit antenna index. The feature vector extraction is conducted as follows:

$$\left| \mathbf{H}_T^n H \mathbf{H}_T^n \right|^2 = \begin{bmatrix} (\mathbf{h}_1^n)^H \\ (\mathbf{h}_2^n)^H \\ \vdots \\ (\mathbf{h}_{N_T}^n)^H \end{bmatrix} \left[ \mathbf{h}_1^n \ \mathbf{h}_2^n \ \dots \ \mathbf{h}_{N_T}^n \right]. \quad (22)$$

To avoid bias, features are normalized using

$$\bar{f}^n(q) = \frac{f^n(q) - E[(f(q))]}{\max(f(q)) - \min(f(q))}, n = 1, \dots, N, q = 1, \dots, N_Q. \quad (23)$$

3. Evaluation of KPI.  
Input data samples are labeled using KPI. FD is considered as the KPI in this work, which is calculated using (14).
4. Class labeling.  
Associate each feature with a label. A transmit antenna subset  $S = \{S_1, S_2, \dots, S_l, \dots, S_{N_s}\}$  is mapped to each feature vector. The feature vector corresponding to each channel matrix  $\hat{\mathbf{H}}_T^n$  is mapped to label  $c^n \in S$ . Hence, the class label vector  $c = [c^1, c^2, \dots, c^N]$ .

### 3.4. ML Methods

This section examines four different supervised ML methods utilized for TAS.

#### 3.4.1. SVM

An SVM model consists of a hyperplane containing representations of distinct classes in multidimensional space [22]. SVM will iteratively create the hyperplane in order to reduce errors. SVM divides datasets into classes in order to determine the best marginal hyperplane. TAS is a multiclass classification problem. The one-versus-one and one-versus-all (OVA) approaches are two common ways of creating multiclass SVM classifiers. This work looks at how OVA can be used to overcome the drawback of complexity in FD-TAS. It creates  $r$  binary SVMs for an  $R$ -class problem. Specifically, for TAS-assisted FGSM,  $R$  SVM models are generated as follows:

- To address this two-class classification problem for the attribute label combinations  $(f^n, c^n)$ , where  $n=1, \dots, N, r \in \{1, 2, \dots, R\}$ , an  $R$  order SVM is constructed. In this case,  $f^n$  represents the  $n$ th row of the normalized attribute matrix  $\hat{\mathbf{H}}_T$ . Samples belonging to the  $r$ th class are characterized by positive labels, while those belonging to the other classes are characterized by negative labels.

- The following approach is used for solving the optimization problem for the two-class SVM:

$$\begin{aligned} & \min_{\boldsymbol{\psi}^r, z^r, \lambda^r} \frac{1}{2} (\boldsymbol{\psi}^r)^T \boldsymbol{\psi}^r + D^r \sum_{n=1}^N \lambda_n^r \\ \text{s.t. } & (\boldsymbol{\psi}^r)^T \phi(\mathbf{f}^n) + z^r \geq 1 - \lambda_n^r, \text{ if } c^n = r, \\ & (\boldsymbol{\psi}^r)^T \phi(\mathbf{f}^n) + z^r \leq -1 + \lambda_n^r, \text{ if } c^n \neq r, \\ & \lambda_n^r \geq 0, r \in \{1, 2, \dots, R\}, \end{aligned} \tag{24}$$

here,  $z^r$  and  $\boldsymbol{\psi}^r$  are linear parameters,  $D^r$  and  $\lambda_n^r$  are the regularization and penalty parameters, respectively, for  $r$ th SVM. By applying kernel function  $\phi(\cdot)$  to an attribute vector in (23), the SVM maps them into a higher-dimensional space.

In this work, the radial basis function (RBF) kernel is utilized to fit the model, which is specified by [22]

$$\phi(\mathbf{f}^i, \mathbf{f}^j) = e^{-\frac{\|\mathbf{f}^i - \mathbf{f}^j\|^2}{2\sigma^2}}. \tag{25}$$

By calculating parameters  $\boldsymbol{\psi}^r$  and  $z^r$  for all valid  $r \in \{1, 2, \dots, R\}$ , the prediction functions obtained can be given as follows:

$$\begin{cases} (\boldsymbol{\psi}^1)^T \phi(\mathbf{f}^{new}) + z^1. \\ \vdots \\ (\boldsymbol{\psi}^r)^T \phi(\mathbf{f}^{new}) + z^r. \end{cases} \tag{26}$$

For a new observation  $\mathbf{H}_T^{new}$ , we utilize (17) to retrieve its feature vector as  $\mathbf{f}^{new}$ , and (26) to obtain its label.

$$I^{SVM} = \arg \max_{r \in \{1, 2, \dots, R\}} \left( (\boldsymbol{\psi}^r)^T \phi(\mathbf{f}^{new}) + z^r \right). \tag{27}$$

### 3.4.2. NB

An NB classifier is a supervised ML technique that uses the Bayesian theorem to solve classification problems [21,27]. Consider the  $N$  features  $[\mathbf{f}^1, \dots, \mathbf{f}^N]$  and  $N_s$  feasible classes in a classification task. One of the  $N_s$  classes must be assigned to the new instance  $\mathbf{H}_T^{new}$ .  $N_s$ -conditional probabilities  $p(l|\mathbf{f}^{new})$ ,  $l=1, \dots, N_s$  are computed by the NB approach. The calculation of joint probability becomes simpler if all features are independent. According to Bayes' theorem, conditional probability is calculated as follows:

$$p(l|\mathbf{f}^{new}) = \frac{p(l)p(\mathbf{f}^{new}|l)}{p(\mathbf{f}^{new})}. \tag{28}$$

In (28), prior probability of the  $l$ th class is given by  $p(l)$ . Here,  $p(\mathbf{f}^{new}|l)$  indicates the likelihood probability and  $p(\mathbf{f}^{new})$  represents evidence probability. The probability of evidence is the same for all classes. As a result, this term can be discarded. Using this method, the class with the greatest probability is selected.

$$I^{NB} = \arg \max_{l \in \{1, 2, \dots, N_s\}} p(l)p(\mathbf{f}^{new}|l). \tag{29}$$

Equation (29) can also be written as follows:

$$I^{NB} = \arg \max_{l \in \{1, 2, \dots, N_s\}} p(l) \prod_{q=1}^{N_Q} p(\mathbf{f}^{new}_q|l). \tag{30}$$

Here,  $\mathbf{f}_q^{new}$  is the  $q$ th feature of the new observation  $\mathbf{f}^{new}$ .

### 3.4.3. DT

Tree-structured classifiers can classify a dataset using internal nodes as attributes, branches as decision rules, and leaves as the inference [21,28]. Two nodes exist in a decision tree—the leaf node and the decision node. Decision nodes make decisions and have multiple branches, whereas leaf nodes generate the output of that decision process and do not contain any additional branches. In a decision tree, an initial question is asked, and based on the answer (yes/no), subtrees are constructed. The attribute value with the highest information gain will be chosen for further branching. The probability  $p_l$  of a sample being owned by a specific class  $l$  is computed using

$$p_l = \frac{\text{freq}(l, T)}{|T|}, \tag{31}$$

here,  $|T|$  is the number of occurrences in  $T$ .  $\text{freq}(l, T)$  represents the number of samples owned by a particular class  $l$ . An entropy of  $T$  is given as

$$EN(T) = - \sum_{l=1}^{N_s} p_l \log_2 p_l. \tag{32}$$

The dataset  $|T|$  is split into  $X$  partitions depending on the domain values of a non-class attribute  $G_l$ . The entropy associated with this is determined using

$$EN(G_l, T) = - \sum_{k=1}^X \frac{|T_k|}{|T|} EN(T). \tag{33}$$

For attribute  $G_l$ , information gain is computed using

$$IFG(G_l) = EN(T) - EN(G_l, T). \tag{34}$$

The information gain is computed for every feature–value combination. When splitting the root node, the feature–value combination that yields the highest information gain is chosen. The process is repeated at all decision nodes. Repeat this process until the maximum depth has been reached. Since the gain will not rise beyond a certain depth, the depth of the DT is regarded as a hyperparameter in the learning process.

### 3.4.4. KNN

This algorithm is lazy; it does not immediately begin learning from the data [22]. Instead, it saves it and performs the action when it is time to classify. For a new instance  $H_T^{new}$ , its feature vector  $f^{new}$  is acquired. When this feature vector is normalized, we obtain  $f^{new}$ . The KNN classifier finds the  $K$ -closest data points amidst  $N$  data points. The KNN classifier performs the following main steps:

Step 1: Determine  $K$  through  $k$ -fold cross-validation for which the model has low variance and reasonably good accuracy.

Step 2: Calculate the FD of a new observation  $f^{new}$  for each data point in the dataset.

$$d(f^n, f^{new}) = \|f^n - f^{new}\|^2. \tag{35}$$

Step 3: Find out the  $K$ -nearest neighbors.

Step 4: For each  $K$ -nearest neighbor, count the data samples.

Step 5: Allocate the test data sample  $H_T^{new}$  to the class label, which obtains the maximum number of votes ( $I^{KNN}$ ).

Despite its apparent simplicity, the cost of computing the FD between data samples and the time complexity increases with the dataset size. One of the disadvantages of KNN is that it works well for small datasets but becomes more difficult and time-expensive with

larger datasets. Identifying the value of  $K$  is always necessary, although it might be difficult at times.

#### 4. Complexity Analysis

Table 4 shows the complexity analysis for traditional FD-TAS and distinct supervised TAS systems.  $\mathbf{H}_T(\chi_i - \chi_j)$  requires  $N_{rx} \times N_T$  computations.  $(\mathbf{H}_T(\chi_i - \chi_j))^H (\mathbf{H}_T(\chi_i - \chi_j))$  needs  $(N_{rx} \times N_T)^2$  computations. These computations are carried out only for one value of FD. This process is repeated  $(N_s 2^{\alpha_{FGSM}} (2^{\alpha_{FGSM}} - 1))$  times for all possible antenna subsets. It is observed that the number of computations required for TAS using the traditional method is tedious. For ML techniques, it depends on the attribute vector length  $N_Q$ . The number of computations required is much less compared to the traditional FD-TAS for a higher-order configuration, i.e., for a larger number of  $N_T$ ,  $N_{rx}$ , and  $\alpha_{FGSM}$ . KNN is the least computationally efficient of all ML approaches because it stores the training data. This analysis is conducted for  $B = 2 \times 10^5$  bits. By decreasing the length of the attribute vector, we can reduce the complexity of ML-based TAS practices. Traditional FD-TAS becomes increasingly complex as these components rise, necessitating the use of ML techniques for TAS to simplify the system.

**Table 4.** Comparison of distinct TAS techniques based on their complexity order.

ML-Based TAS Scheme	Order of Complexity	Configuration 1 ( $N_{tx} = N_{rx} = 4$ , $N_T = 5, M = 4$ )	Configuration 2 ( $N_{tx} = N_{rx} = 6$ , $N_T = 7, M = 16$ )
FD-TAS	$\mathcal{O}(N_s N_T^2 N_{rx}^2 2^{\alpha_{FGSM}} (2^{\alpha_{FGSM}} - 1))$ [20]	$1.98 \times 10^6$	$3.23 \times 10^9$
SVM-based TAS	$\mathcal{O}(N_Q^2 + N_Q)$ [22]	420	1806
NB-based TAS	$\mathcal{O}(N_Q)$ [21,27]	20	42
DT-based TAS	$\mathcal{O}(N_Q \log_2 N_Q)$ [21,28]	86.44	226.48
KNN-based TAS	$\mathcal{O}(BN_Q)$ [22]	$4 \times 10^6$	$8.4 \times 10^6$

#### 5. Discussion on the Simulation Results

This section presents an ABER analysis of the classic FD-TAS method, four distinct ML-based TAS methods, and the FGSM-NTAS scheme. Table 5 lists the parameters used in this investigation. In MATLAB 2022b, all ML models are trained using the Statistics and ML toolbox. All the simulation results are generated for distinct values of  $N_T$ ,  $N_{tx}$ ,  $N_{rx}$ , and  $M$ . FD-TAS and ML-based TAS are labeled as  $(N_T, N_{tx}, N_{rx}, M)$  in the figures, while FGSM-NTAS is labeled as  $(N_{tx}, N_{rx}, M)$ .

**Table 5.** Simulation parameters.

Parameters	Values
$N_T$	4, 5, 6
$N_{tx}$	3, 4
$N_{rx}$	4
$M$	4, 16
Constellation mapping	Quadrature amplitude modulation (QAM)
$\alpha_{FGSM}$	4, 5, 6
$B$	$2 \times 10^5$
Fading environment	Uncorrelated Rayleigh flat fading

From Figure 4, it is observed that with an increased dataset size, ML models show an improvement in classification accuracy. Since the accuracy of classification does not improve any further after 10,000 samples, all models are trained for the same number of samples. Table 6 compares the classification accuracies of several ML algorithms with different sample counts. SVM has the highest accuracy in classification (71.1%), NB has the second greatest accuracy (70.2%), followed by KNN (67.2%), and DT has the lowest classification accuracy (43.8%).

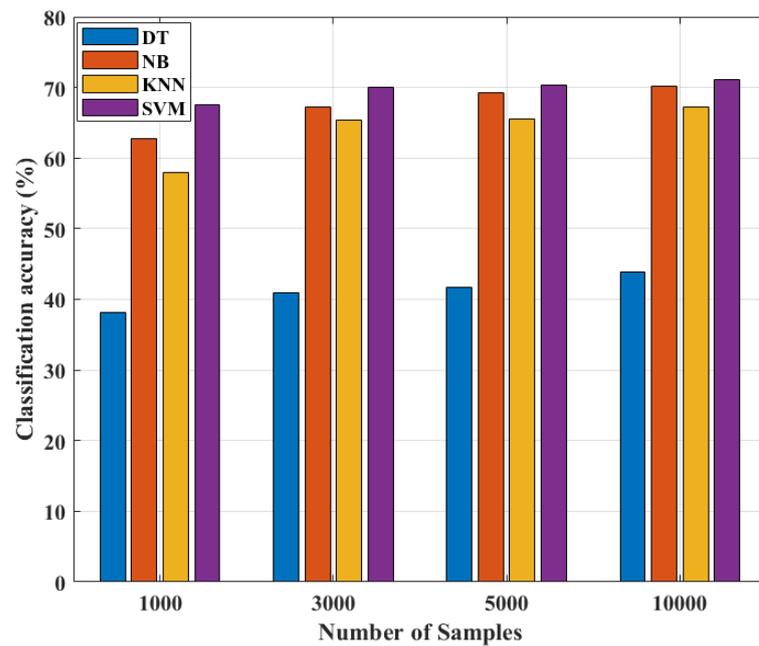


Figure 4. Comparing the classification accuracy vs. dataset size for several ML algorithms.

Table 6. Comparing the classification accuracy vs. dataset size for several ML algorithms.

Number of Samples	Achieved Percentage of Accuracy			
	DT	NB	KNN	SVM
1000	38.2%	62.8%	57.9%	67.6%
3000	40.9%	67.2%	65.3%	70%
5000	41.7%	69.2%	65.6%	70.3%
10,000	43.8%	70.2%	67.2%	71.1%

Cross-validation is an approach that allows the model to learn from many train–test splits. This provides a better idea of how well the trained model will perform on data that have not been seen before. In contrast, hold-out is based on a single train–test split. As a result, the hold-out method’s score is influenced by the way the data are divided into train and test sets. Figure 5 and Table 7 show how altering the value of  $k$ -folds during cross-validation improves classification accuracy. The highest classification accuracy is observed for  $k = 10$ . As a result, the value of  $k$  is set to 10.

Table 7. Comparing the distinct TAS techniques based on their accuracy performances.

ML-Based TAS Scheme	Achieved Percentage of Accuracy		
	$k = 3$	$k = 5$	$k = 10$
SVM-based TAS [22]	70.1%	70.7%	71.1%
NB-based TAS [21,27]	69.5%	69.9%	70.2%
DT-based TAS [21,28]	42.9%	43.3%	43.8%
KNN-based TAS [22]	66.2%	66.8%	67.2%

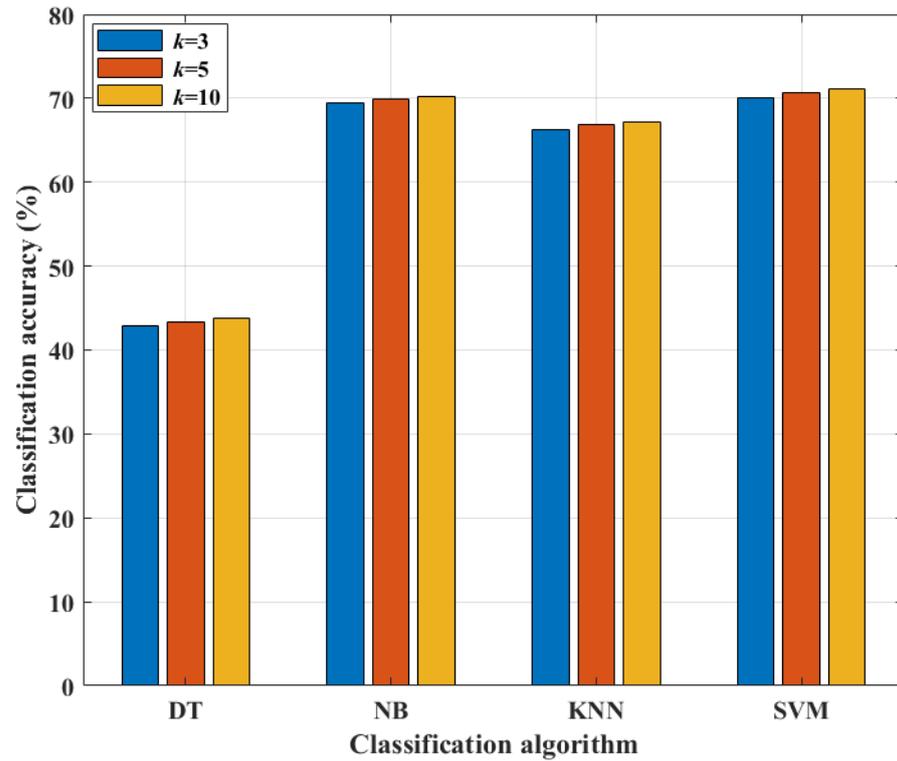


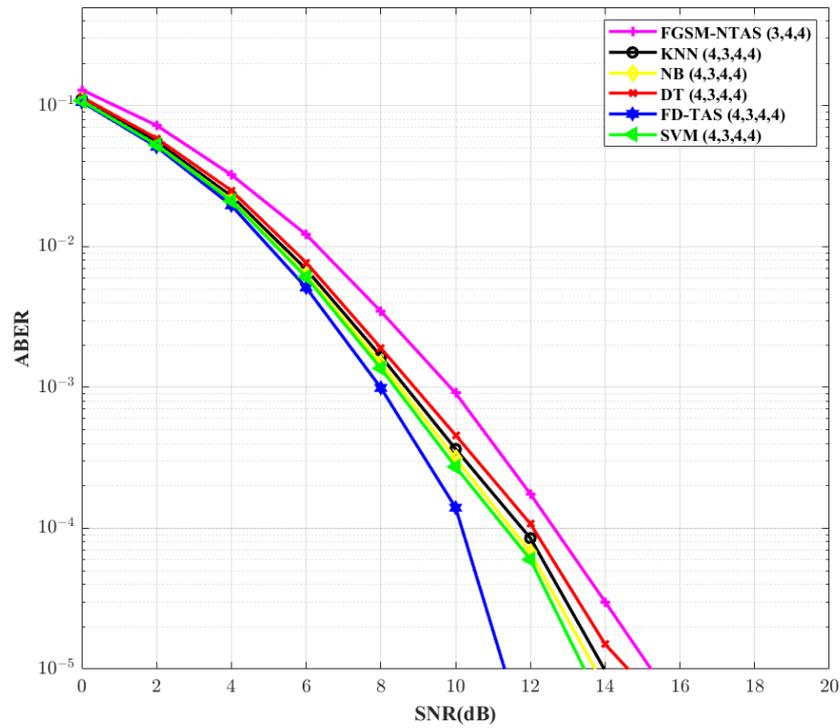
Figure 5. Comparing the classification accuracies of several ML algorithms for distinct values of  $k$ .

After performing normalization, these models are trained for four distinct features, as discussed earlier in subsection 3.3. The complexities of all features based on their attribute vector lengths are listed in Table 8. The worst-performing attribute is  $\angle(h_{i,j}^n)$ . Attribute 4 performs well only for DT, but its complexity increases with a higher-order configuration, so it is discarded. Combining features lengthens the feature vector and increases model complexity; hence, feature combination is avoided in this study. The absolute feature outperforms the other features in SVM, NB, and KNN, and its complexity is low; thus, only the absolute feature is utilized for this work.

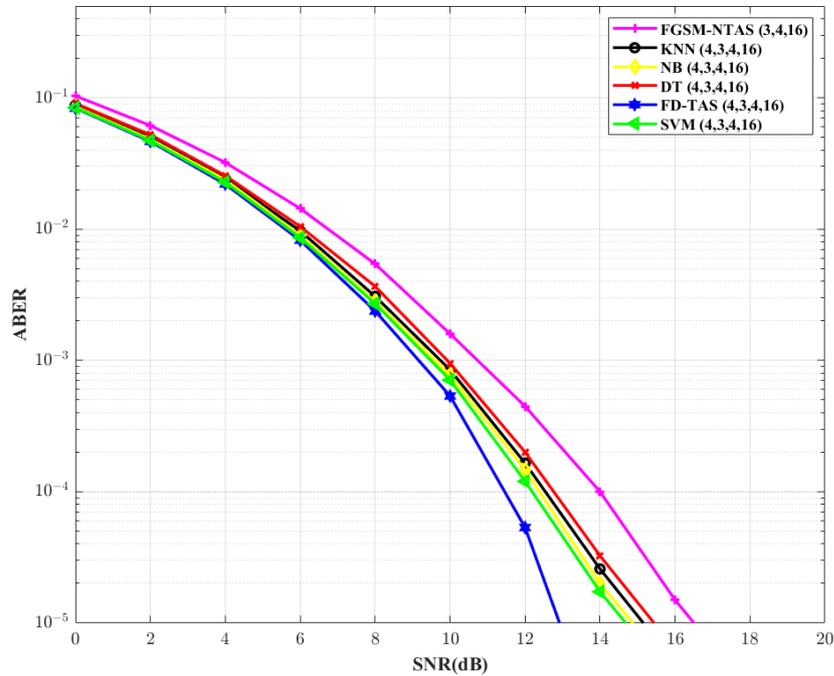
Table 8. Complexity associated with different features based on the feature vector length.

Selected Feature	Length of the Feature Vector ( $N_Q$ )	Configuration 1 ( $N_{rx} = 2, N_T = 6$ )	Configuration 2 ( $N_{rx} = 4, N_T = 10$ )
$ h_{i,j}^n $	$N_{rx} \times N_T$	12	40
$\angle(h_{i,j}^n)$	$N_{rx} \times N_T$	12	40
$ h_{i,j}^n ^2$	$N_{rx} \times N_T$	12	40
$ h_{i,j}^{nH} h_{i,j}^n ^2$	$N_T \times N_T$	36	100

After conventional FD-TAS, the SVM-based TAS scheme outshines other ML-based systems, as shown in Figure 6. In comparison to the FGSM-NTAS method, the SVM-based TAS scheme cut downs the SNR need by  $\sim 2$  dB. This analysis is conducted for  $N_T = 4$ ,  $N_{tx} = 3$ ,  $N_{rx} = 4$ , and  $M = 4$ . The ABER performance of the system is re-examined in Figure 7 by increasing the constellation order to 16. For larger values of  $M$ , the SNR requirement for all ML-based methods increases. For FGSM-NTAS and conventional FD-TAS SNR gain requirements increase by  $\sim 1.3$  dB and  $\sim 1.61$  dB, respectively. SVM outperforms the FGSM-NTAS system in this circumstance as well, with a gain improvement of  $\sim 2$  dB.



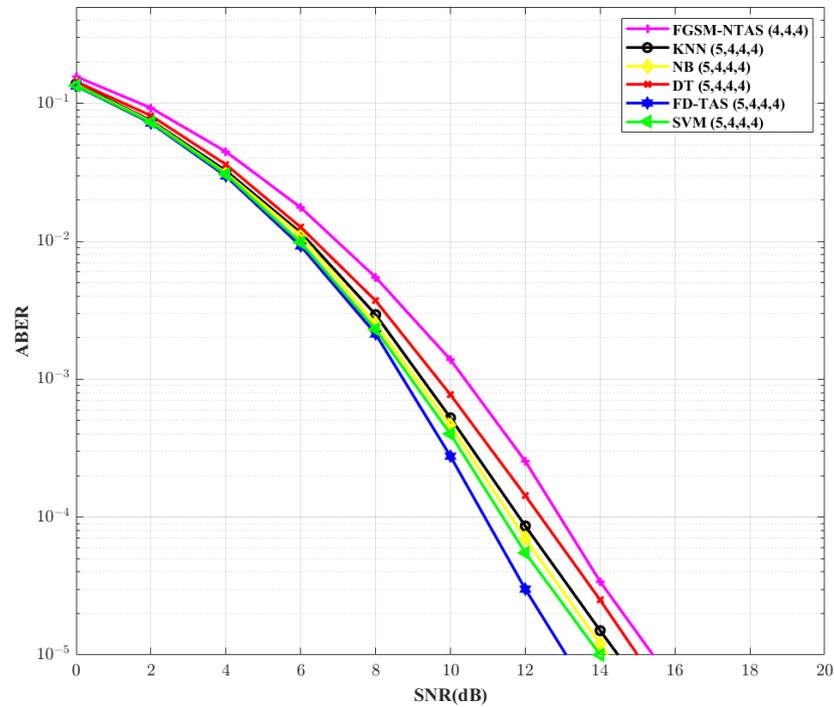
**Figure 6.** Comparison between ABER and SNR of FGSM for distinct TAS practices ( $N_T = 4$ ,  $N_{tx} = 3$ ,  $N_{rx} = 4$ , and  $M = 4$ ).



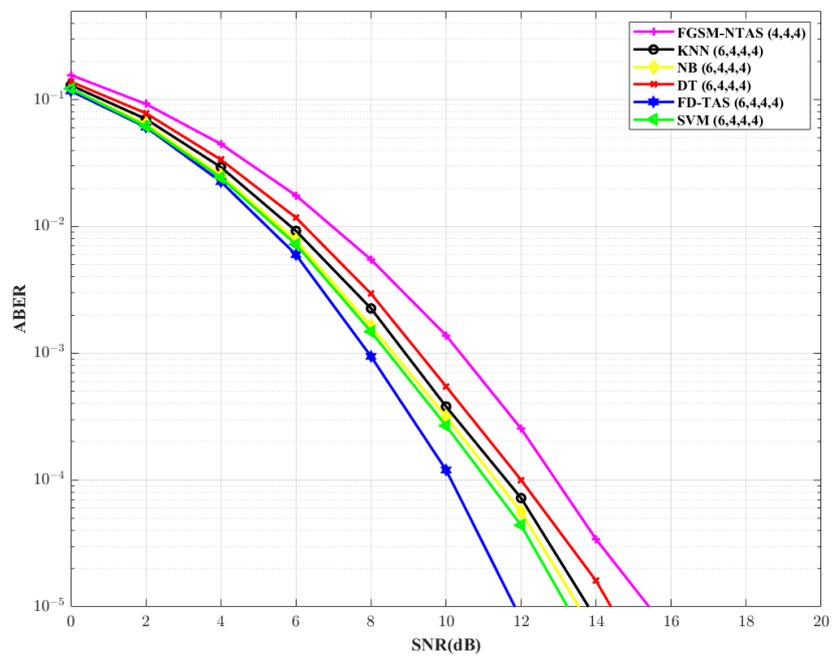
**Figure 7.** Comparison between ABER and SNR of FGSM for distinct TAS practices ( $N_T = 4$ ,  $N_{tx} = 3$ ,  $N_{rx} = 4$ , and  $M = 16$ ).

Figure 8 analyzes the ABERs of different TAS schemes for  $N_T = 5$ ,  $N_{tx} = 4$ ,  $N_{rx} = 4$ , and  $M = 4$  configuration. It is found that SVM-based TAS outperforms FGSM-NTAS by  $\sim 1.42$  dB in terms of SNR gain. By increasing  $N_T$  to 6, the analysis is repeated in Figure 9. For traditional FD-TAS, the required SNR drops by  $\sim 1.28$  dB. For SVM, NB, KNN, and DT-based TAS systems, respectively, there is a reduction of  $\sim 0.78$  dB,  $\sim 0.66$  dB,  $\sim 0.68$  dB,

and  $\sim 0.59$  dB in the SNR requirements. In comparison to the FGSM-NTAS system, the SVM-based TAS technique gains  $\sim 2.2$  dB. Table 9 lists the SNR gain achieved by the proposed ML-based TAS practices over traditional FGSM-NTAS.



**Figure 8.** Comparison between ABER and SNR of FGSM for distinct TAS practices ( $N_T = 5$ ,  $N_{tx} = 4$ ,  $N_{rx} = 4$ , and  $M = 4$ ).



**Figure 9.** Comparison between ABER and SNR of FGSM for distinct TAS practices ( $N_T = 6$ ,  $N_{tx} = 4$ ,  $N_{rx} = 4$ , and  $M = 4$ ).

**Table 9.** SNR gain achieved by distinct TAS practices over FGSM-NTAS.

ML-Based TAS Scheme	Improvement in SNR gain (dB)			
	Configuration 1 ( $N_T = 4, N_{tx} = 3,$ $N_{rx} = 4,$ and $M = 4$ )	Configuration 2 ( $N_T = 4, N_{tx} = 3,$ $N_{rx} = 4,$ and $M = 16$ )	Configuration 3 ( $N_T = 5, N_{tx} = 4,$ $N_{rx} = 4,$ and $M = 4$ )	Configuration 4 ( $N_T = 6, N_{tx} = 4,$ $N_{rx} = 4,$ and $M = 4$ )
SVM-based TAS [22]	1.76	1.89	1.44	2.2
NB-based TAS [21,27]	1.49	1.65	1.21	1.86
DT-based TAS [21,28]	0.59	1.08	0.46	1.03
KNN-based TAS [22]	1.27	1.39	0.96	1.61

## 6. Conclusions

Four distinct ML-based TAS techniques are suggested and implemented for a small-scale MIMO-based FGSM system in this work. For different values of  $N_T$ ,  $N_{tx}$ , and  $M$ , the recommended algorithm's ABER performance is compared to those of FD-TAS and FGSM-NTAS. The SVM-based TAS scheme surpasses all proposed ML-based TAS approaches in the simulations. The ABER performance of SVM is poorer than FD-TAS but the SVM-based TAS scheme is more computationally efficient than FD-TAS. SVM-based TAS attains the highest classification accuracy ( $\sim 71.1\%$ ) and a minimal SNR gain of  $\sim 2.2$  dB compared to FGSM-NTAS. As a result, it may be a viable option for future-generation networks. These schemes, as well as DL-based architectures, can be implemented in the future for advanced SM variants. The proposed TAS strategies could be extended to massive MIMO systems.

**Author Contributions:** All authors contributed to the manuscript. Conceptualization: H.K.J. and V.B.K.; article gathering and sorting: H.K.J., V.B.K., A.M., A.L.I., P.S. and A.C.; resources: H.K.J., V.B.K., A.M., A.L.I., P.S. and A.C.; supervision: V.B.K., A.M., A.L.I., P.S. and A.C.; validation: V.B.K., A.M., A.L.I., P.S. and A.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work of Agbotiname Lucky Imoize is supported in part by the Nigerian Petroleum Technology Development Fund (PTDF) and in part by the German Academic Exchange Service (DAAD) through the Nigerian-German Postgraduate Program under Grant 57473408.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

5G	Fifth-Generation
6G	Sixth-Generation
ABER	Average Bit Error Rate
AWGN	Additive White Gaussian Noise
DL	Deep Learning
DT	Decision Tree
EE	Energy Efficiency
FD	Free Distance

FGSM	Fully Generalized Spatial Modulation
IM	Index Modulation
KNN	K-Nearest Neighbor
MIMO	Multiple-Input Multiple-Output
ML	Machine Learning
NB	Naïve Bayes
QAM	Quadrature Amplitude Modulation
SE	Spectral Efficiency
SM	Spatial Modulation
SNR	Signal-to-Noise Ratio
SVM	Support Vector Machine
TAS	Transmit Antenna Selection
UE	User Equipment

## References

1. Turker, I.; Tan, S.O. Machine Learning vs. Deep Learning in 5G Networks—A Comparison of Scientific Impact. *arXiv* **2022**, arXiv:2210.07327.
2. De Alwis, C.; Kalla, A.; Pham, Q.V.; Kumar, P.; Dev, K.; Hwang, W.J.; Liyanage, M. Survey on 6G frontiers: Trends, applications, requirements, technologies and future research. *IEEE Open J. Commun. Soc.* **2021**, *2*, 836–886.
3. Jadhav, H.K.; Kumaravelu, V.B. Deep Learning-Assisted Transmit Antenna Classifiers for Fully Generalized Spatial Modulation: Online Efficiency Replaces Offline Complexity. *Appl. Sci.* **2023**, *13*, 5134. [[CrossRef](#)]
4. Lee, H.; Lee, B.; Yang, H.; Kim, J.; Kim, S.; Shin, W.; Shim, B.; Poor, H.V. Towards 6G hyper-connectivity: Vision, challenges, and key enabling technologies. *J. Commun. Networks* **2023**, *25*, 344–354. [[CrossRef](#)]
5. Shahjalal, M.; Kim, W.; Khalid, W.; Moon, S.; Khan, M.; Liu, S.; Lim, S.; Kim, E.; Yun, D.W.; Lee, J.; et al. Enabling technologies for AI empowered 6G massive radio access networks. *ICT Express* **2023**, *9*, 341–355.
6. Wang, Z.; Du, Y.; Wei, K.; Han, K.; Xu, X.; Wei, G.; Tong, W.; Zhu, P.; Ma, J.; Wang, J.; et al. Vision, application scenarios, and key technology trends for 6G mobile communications. *Sci. China Inf. Sci.* **2022**, *65*, 151301.
7. Chowdhury, M.Z.; Shahjalal, M.; Ahmed, S.; Jang, Y.M. 6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions. *IEEE Open J. Commun. Soc.* **2020**, *1*, 957–975.
8. Chochliouros, I.P.; Kourtis, M.A.; Spiliopoulou, A.S.; Lazaridis, P.; Zaharis, Z.; Zarakovitis, C.; Kourtis, A. Energy efficiency concerns and trends in future 5G network infrastructures. *Energies* **2021**, *14*, 5392.
9. Mesleh, R.Y.; Haas, H.; Sinanovic, S.; Ahn, C.W.; Yun, S. Spatial modulation. *IEEE Trans. Veh. Technol.* **2008**, *57*, 2228–2241. [[CrossRef](#)]
10. Basar, E.; Wen, M.; Mesleh, R.; Di Renzo, M.; Xiao, Y.; Haas, H. Index modulation techniques for next-generation wireless networks. *IEEE Access* **2017**, *5*, 16693–16746.
11. Castillo-Soria, F.R.; Cortez-González, J.; Ramirez-Gutierrez, R.; Maciel-Barboza, F.M.; Soriano-Equigua, L. Generalized quadrature spatial modulation scheme using antenna grouping. *ETRI J.* **2017**, *39*, 707–717.
12. Asaad, S.; Rabiei, A.M.; Müller, R.R. Massive MIMO with antenna selection: Fundamental limits and applications. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 8502–8516. [[CrossRef](#)]
13. Gudla, V.V.; Kumaravelu, V.B.; Murugadass, A. Transmit antenna selection strategies for spectrally efficient spatial modulation techniques. *Int. J. Commun. Syst.* **2022**, *35*, e5099. [[CrossRef](#)]
14. Rajashekar, R.; Hari, K.; Hanzo, L. Antenna selection in spatial modulation systems. *IEEE Commun. Lett.* **2013**, *17*, 521–524. [[CrossRef](#)]
15. Pillay, N.; Xu, H. Comments on “Antenna Selection in Spatial Modulation Systems”. *IEEE Commun. Lett.* **2013**, *17*, 1681–1683. [[CrossRef](#)]
16. Zhou, Z.; Ge, N.; Lin, X. Reduced-complexity antenna selection schemes in spatial modulation. *IEEE Commun. Lett.* **2013**, *18*, 14–17. [[CrossRef](#)]
17. Pillay, N.; Xu, H. Low-complexity transmit antenna selection schemes for spatial modulation. *IET Commun.* **2014**, *9*, 239–248. [[CrossRef](#)]
18. Junior, E.N.; Theis, G.; Santos, E.L.d.; Mariano, A.A.; Brante, G.; Souza, R.D.; Taris, T. Energy Efficiency Analysis of MIMO Wideband RF Front-End Receivers. *Sensors* **2020**, *20*, 7070. [[CrossRef](#)]
19. Bereyhi, A.; Asaad, S.; Mueller, R.R. Stepwise transmit antenna selection in downlink massive multiuser MIMO. In Proceedings of the WSA 2018; 22nd International ITG Workshop on Smart Antennas, VDE, Bochum, Germany, 14–16 March 2018; pp. 1–8.
20. Yang, P.; Xiao, Y.; Li, L.; Tang, Q.; Yu, Y.; Li, S. Link adaptation for spatial modulation with limited feedback. *IEEE Trans. Veh. Technol.* **2012**, *61*, 3808–3813.
21. Jadhav, H.K.; Kumaravelu, V.B. Transmit antenna selection for spatial modulation based on machine learning. *Phys. Commun.* **2022**, *55*, 101904. [[CrossRef](#)]
22. Yang, P.; Zhu, J.; Xiao, Y.; Chen, Z. Antenna selection for MIMO system based on pattern recognition. *Digit. Commun. Networks* **2019**, *5*, 34–39.

23. He, D.; Liu, C.; Quek, T.Q.; Wang, H. Transmit antenna selection in MIMO wiretap channels: A machine learning approach. *IEEE Wirel. Commun. Lett.* **2018**, *7*, 634–637.
24. Yang, P.; Xiao, Y.; Xiao, M.; Guan, Y.L.; Li, S.; Xiang, W. Adaptive spatial modulation MIMO based on machine learning. *IEEE J. Sel. Areas Commun.* **2019**, *37*, 2117–2131. [[CrossRef](#)]
25. Liu, H.; Xiao, Y.; Yang, P.; Fu, J.; Li, S.; Xiang, W. Transmit Antenna Selection for Full-Duplex Spatial Modulation Based on Machine Learning. *IEEE Trans. Veh. Technol.* **2021**, *70*, 10695–10708.
26. Altın, G.; Arslan, İ.A. Joint transmit and receive antenna selection for spatial modulation systems using deep learning. *IEEE Commun. Lett.* **2022**, *26*, 2077–2080.
27. Shi, Y.; Lu, X.; Niu, Y.; Li, Y. Efficient jamming identification in wireless communication: Using small sample data driven naive bayes classifier. *IEEE Wirel. Commun. Lett.* **2021**, *10*, 1375–1379. [[CrossRef](#)]
28. León, J.P.A.; de la Cruz Llopis, L.J.; Rico-Novella, F.J. A Machine Learning Based Distributed Congestion Control Protocol for Multi-Hop Wireless Networks. *Comput. Netw.* **2023**, *231*, 109813.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.