

การตัดสินใจการกลับบิตด้วยเทคนิคการเรียนรู้ของเครื่องในระบบการบันทึกเชิงแม่เหล็ก

Bit Flipping Decision based on Machine Learning Technique in Magnetic Recording System

Suphakit Chareonsit¹ Anucha Tungkasthan² Wiparat Busyatras³ and Pitaya Poompuang⁴

Faculty of Science and Technology Rajamangala University of Technology Thanyaburi, Pathum Thani, Thailand

suphakit_c@mail.rmutt.ac.th¹, anucha_t@rmutt.ac.th², wiparat_b@rmutt.ac.th³, p_pitaya@rmutt.ac.th⁴

บทคัดย่อ

บทความนี้ศึกษาปัจจัยหลักและวิธีการในการทำนายการกลับบิต เทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors) ถูกนำเสนอสำหรับการสร้างแบบจำลองการกลับบิตของระบบบีทีเอ็มอาร์เพื่อเพิ่มความหนาแน่นของพื้นที่สูงในระบบบันทึกด้วยแม่เหล็ก ระบบบีทีเอ็มอาร์มีการรบกวนหลายอย่างเพื่อลดประสิทธิภาพ เช่น การแทรกสอดสองมิติหรือการรบกวนระหว่างแทร็ก (ITI) และการรบกวนระหว่างสัญลักษณ์ (ISI) สัญลักษณ์รบกวนจากตำแหน่งประสิทธิภาพของอัตราบิตผิดพลาด (BER) สามารถปรับปรุงได้กับแต่ละกระบวนการของระบบบีทีเอ็มอาร์ เช่น การเข้ารหัส การถอดรหัส งานวิจัยชิ้นนี้แสดงให้เห็นถึงข้อดีของการตัดสินใจกลับบิตขึ้นอยู่กับความสัมพันธ์ของค่าที่อธิบายโดยอัตราส่วนความน่าจะเป็นของบันทึกก่อนหน้า และใช้เทคนิคการกลับบิตเพื่อลดผลกระทบของการแทรกสอดระหว่างแทร็ก (ITI) และ สัญลักษณ์แทรกแซงระหว่างสัญลักษณ์ (ISI) โดยใช้ค่าอัตราส่วนที่ดีที่สุดสำหรับการตัดสินใจกลับบิต ค่าเกณฑ์ได้มาจากการทดสอบระบบด้วยประสิทธิภาพอัตราความผิดพลาดบิต (BER) ที่ดีที่สุด ประสิทธิภาพของวิธีการที่เสนอนี้ให้ประสิทธิภาพสูง ดังนั้น งานวิจัยนี้จึงต้องการเน้นที่ อัลกอริทึมเคเนียร์เรสเนเบอร์แสดงถึงระยะทางแบบยุคลิดสำหรับการ

คำนวณการทำนายแบบจำลอง ในการเตรียมข้อมูล เราเริ่มศึกษาค่าที่สำคัญสำหรับการเตรียมชุดข้อมูลเรียนรู้ของค่าสัญญาณอ่านกลับ ,ค่าความน่าจะเป็นความสามารถของค่าสัญญาณอ่านกลับ, ค่าเฉลี่ยของค่าสัญญาณอ่านกลับ และผลรวมของค่าสัญญาณอ่านกลับ เราพบว่าเปอร์เซ็นต์ความแม่นยำของอัลกอริทึมเคเนียร์เรสเนเบอร์ให้ความแม่นยำในการทำนายที่สูงกว่า 90% ในทุกระดับสัญญาณรบกวน ผลลัพธ์นี้แสดงให้เห็นถึงความสำคัญของการใช้เทคนิควิทยาศาสตร์ข้อมูลเพื่อเพิ่มประสิทธิภาพการประมวลผลทั้งหมดในระบบบีทีเอ็มอาร์ด้วยระบบการเข้ารหัสมอดูเลชันแบบสองมิติ

คำสำคัญ: โมเดลการกลับบิต, การแทรกสอดสองมิติ, เคเนียร์เรสเนเบอร์ (K-Neighbors Classifier), การเรียนรู้ของเครื่อง Machine learning (ML), เทคโนโลยีการบันทึกข้อมูลแบบบิตแพทเทิร์นมีเดีย Bit-patterned media recording (BPMR)

Abstract

This paper studies the key factors and methods that make up the prediction of bit inversion. The K-Nearest Neighbors technique is proposed to model the bit inversion of the BPMR

system to increase high space density in magnetic recording systems. The BPMR system has several interferences to reduce its performance such as two-dimensional interference or inter-track interference (ITI) and inter-symbol interference (ISI). Bit error rate (BER) performance can be improved with individual processes of BPMR systems such as encoding, and decoding. That is explained by the likelihood ratio of the previous record and use bit inversion techniques to mitigate the effects of (ITI) and (ISI) by using the best ratio for bit inverting decisions. Threshold values were obtained from testing the system with the best bit error rate (BER) performance. The performance of this proposed method is highly efficient. Therefore, this research aims to focus on the K-Nearest Neighbors algorithm that represents Euclidean distance for model prediction calculations. In a data preparation process, we began to study four values for the preparation of a learning data set of read-back signal values, probability read-back signal capability, average read-back signal value, and the sum of the read back signal values. We found that the accuracy percentage of the K-Nearest Neighbors algorithm provided prediction accuracy which was greater than 90% at all noise levels. This result shows the importance of using the data science techniques

to enhance all processing performance in BPMR with two-dimensional modulation encoding.

Keywords: Bit flipping model, 2D interference, K-Nearest Neighbors (KNN), Machine Learning (ML), Bit-pattern media recording (BPMR).

1. Introduction

The bit patterned media recording (BPMR) system is a technology that increases the high areal density in magnetic recording systems. Several interferences decrease the performance of the BPMR system, such as 2D interference or Inter-Track Interference (ITI) and Inter-Symbol Interference (ISI), Track mis-registration (TMR), and position jitter noise.

Performance can be decreased in each process of the BPMR system, such as the encoding, decoding, and decision of the record bit method processes. The prior work decreased the ITI effect of the system by using a novel soft-information flipping scheme to describe the bit-flipping technique based on a priori log-likelihood ratios of the rate-5/6 encoding constraint in the BPMR system. The research shows that criteria of decision making can gain the performance to flip a bit. However, the process of this bit flipping depends on only a value relation explained by the prior log-likelihood ratios. The bit flipping technique was used to decrease the effect of ITI and ISI by

using the best ratio value for bit flipping decisions. The threshold value was achieved from system testing with the best bit error rate (BER) performance. The performance of this proposed method offers high-performance gain. It is very interesting for some significance of this flipping technique in the final process of BER performance. If the system can optimize the best threshold, it will also achieve high performance.

Nowadays, data science has become a very interesting process used for analyzing data. So, in this paper, we propose applying data science techniques, the K-Nearest Neighbors (KNN) algorithm, as a prediction model to estimate if the bit would be flip or not on the BPMR system with 5/6 coding.

2. Related Works

In the new technology of hard disk drive (HDD), one of high areal density of new technology in the future is bit patterned media recording (BPMR) system represent as the road map of HDD suggestion, heat-assisted magnetic recording (HAMR) will soon replace conventional perpendicular magnetic recording (PMR). The first few generations of HAMR based HDDs are expected to reach 2 Tb/in² by 2019. Around 2025, bit patterned magnetic recording (BPMR) is expected to enter the market, along with the shingled magnetic recording (SMR) and two-

dimensional magnetic recording (TDMR). To realize extremely of high-density approaching, 10 Tb/in², heated-dot magnetic recording, a combination of BPMR and HAMR would be employed [1]. Several research types focus on increases in the BPMR system by using the two-dimensional (2D) Modulation coding method to avoid ITI and ISI in the BPMR system [2-6]. In our previous method, we used the 2D modulation code combined with the bit flipping technique based on ratio value, and this work did not use the data science techniques for finding the appropriate threshold value [7].

Furthermore, some research using the data science techniques for improving the performance of BPMR [8-10] is very interesting and has some significance to focus on the data science technique or machine learning method to find a model to estimate a flipping bit BPMR system in our previous method [7]. Moreover, the 2D modulation coding method can support that well data preparation patterned for using data science techniques to create a model. In the perpendicular recording of hard disk, data has a limitation of recording capacity, areal density (AD) not more than 1 Tb/in². In today's hard disk drive industry, technology is competing to increase the storage space in hard disk drives. Disk drives include Two-dimensional magnetic data recording, two-dimensional magnetic

recording (TDMR), microwave-assisted magnetic recording (MAMR), heat-assisted magnetic recording (HAMR), bit-patterned media recording (BPMR) [8]. We are interested in the BPMR system that solves the problem of recording data at the capacity of areal density (AD) 4 Tb/in². [9]. In [7] The soft output bit inversion technique of the Viterbi algorithm is used to decide the bit inversion using priori log-likelihood ratios using data science techniques. However, the process of this bit flipping depends on only a value of the prior log-likelihood ratios. The bit flipping technique was used to decrease the effecting of ITI and ISI by using the best ratio value for bit flipping decisions without data science technique and performance of this proposed method offers high-performance gain. It is very interesting for some significance in this flipping technique in the final process of BER performance.

Therefore, this research would like to focus on multi-values conditions for predicting the bit flipping by using one of the data science techniques, the K-Nearest Neighbors (KNN) is proposed to create the bit flipping model the BPMR system with 5/6 coding. The KNN algorithm represents the Euclidean distance for calculating the model prediction. In the data preparation, we begin studying the four values for preparing the Data train: the readback signal

values, the Pop ability of readback signal value, the mean of readback signal value, and the summation of readback signal value. Then, the Flip or Not Flip are selected for the prediction classification. Finally, the predictive models are generated through the KNN algorithm. In this simulation, we focus on the binary classification scheme of supervised learning. Binary classification is a function of classifying a set of elements into two groups, meaning that the results in our prediction models can be as few as Flip or Not Flip. We found that the accuracy percentage of the KNN algorithm offers high prediction accuracy of more than 90% in all SNR levels as 2, 5 and 10 dB. The rest of this paper is organized as follows: Section 2 briefly describes the related work, section 3 briefly the proposed method; Section 4 explains the data preparation and the KNN algorithm; Section 5 explains labels prediction. The simulation results are given in Section 6; and finally, Section 7 concludes this paper.

3. Proposed Method

In the data preparation process, we begin to study the readback signal in the BPMR process system [10-14] as shown in Fig1. We estimate the bit flipping by using the output of the 2D SOVA detector: the readback signal values, $r_{l,k}$, the probability of readback signal

value, $p_{l,k}$, the mean of readback signal value, $m_{l,k}$, and the summation of readback signal value, $s_{l,k}$. Then, the two values, Flip or Not Flip, are selected to be the target label for the prediction classification. Finally, we generate the predictive based on using the KNN algorithm.

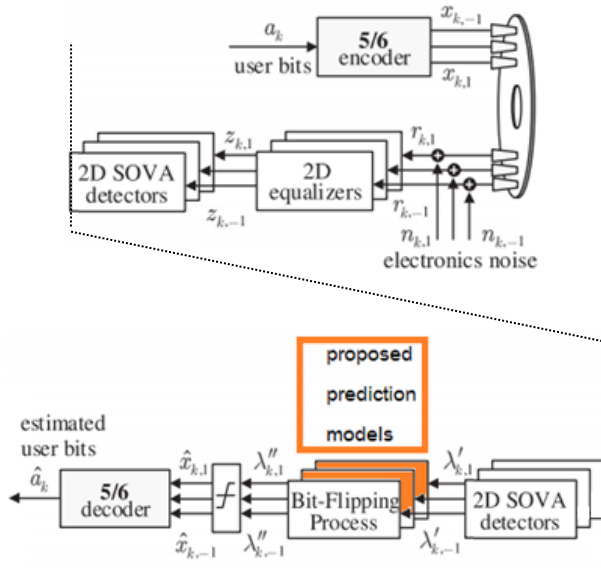


Fig. 1 The BPMPR channel model diagram with the rate 5/6 modulation code scheme [7].

4. Data Preparation

In the data preparation process, the two set of data were prepared as a training set and the testing set for the KNN algorithm. The table1 and the table2 show some examples of data for training set and testing set respectively. Both share the same data structure consisting of the readback signal values, $r_{l,k}$, the probability of readback signal value, $p_{l,k}$, the mean of readback signal value, $m_{l,k}$, and the summation of readback signal value, $s_{l,k}$. As we need to

predict that the bit should be Flip or Not Flip, the label column is included in the structure of both tables.

Table 1. The example of data training preparation for finding models to flip bit.

$r_{l,k}$	$p_{l,k}$	$m_{l,k}$	$s_{l,k}$	Labels
2.2823	0.4160	2.8879	8.6636	Not Flip
0.1822	0.2811	2.1530	6.4589	Not Flip
-1.0650	0.3858	2.7745	8.3235	Not Flip
-0.6414	0.3454	1.2585	3.7754	Not Flip
-0.41762	0.020496	1.0251	3.0752	Flip
-0.30488	0.023442	1.8689	5.6067	Flip
0.037616	0.22123	0.12706	0.38118	Flip
-1.034	0.038474	2.1221	6.3663	Flip

Table 2. The example of testing for finding models to flip bit with label values.

$r_{l,k}$	$p_{l,k}$	$m_{l,k}$	$s_{l,k}$	Labels
2.2632	0.3520	3.1084	9.3252	Not Flip
0.6725	0.2346	1.2523	3.7569	Not Flip
-1.0206	0.1825	1.2706	3.8117	Flip
1.7635	0.17542	1.8005	5.4015	Flip
-1.7933	0.10598	1.1369	3.4107	Flip

5. Label Prediction

In the modelling process, we apply the K-Nearest Neighbors (KNN) model to predict the criteria of the bit flipping in the BPMPR system with 5/6 coding. The KNN algorithm calculates distance of examples using Euclidean equation as (1).

$$D_{Euclidean} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_L - y_L)^2} \quad (1)$$

To test the model, first we remove the values of the label column.as shown in Table3.

Table 3. The example of testing for finding models to flip bit without label values.

$r_{l,k}$	$P_{l,k}$	$m_{l,k}$	$S_{l,k}$	Labels
2.2632	0.3520	3.1084	9.3252	?
0.6725	0.2346	1.2523	3.7569	?
-1.0206	0.1825	1.2706	3.8117	?
1.7635	0.17542	1.8005	5.4015	?
-1.7933	0.10598	1.1369	3.4107	?

$$D_{1,1} = \sqrt{(2.2823 - 2.2632)^2 + (0.4160 - 0.3520)^2 + (2.8879 - 3.1084)^2 + (8.6636 - 9.3252)^2} = 0.701$$

$$D_{2,1} = \sqrt{(0.1822 - 2.2632)^2 + (0.2811 - 0.3520)^2 + (2.153 - 3.1084)^2 + (6.4589 - 9.3252)^2} = 3.669$$

$$D_{3,1} = \sqrt{(-1.065 - 2.2632)^2 + (0.3858 - 0.3520)^2 + (2.7745 - 3.1084)^2 + (8.3235 - 9.3252)^2} = 3.492$$

$$D_{4,1} = \sqrt{(-0.6414 - 2.2632)^2 + (0.3454 - 0.3520)^2 + (1.2585 - 3.1084)^2 + (3.7754 - 9.3252)^2} = 5.996$$

$$D_{5,1} = \sqrt{(-0.4176 - 2.2632)^2 + (0.0204 - 0.352)^2 + (1.0251 - 3.1084)^2 + (3.0752 - 9.3252)^2} = 7.120$$

$$D_{6,1} = \sqrt{(-0.3048 - 2.2632)^2 + (0.0234 - 0.352)^2 + (1.8689 - 3.1084)^2 + (5.6067 - 9.3252)^2} = 4.698$$

$$D_{7,1} = \sqrt{(0.0376 - 2.2632)^2 + (0.22123 - 0.352)^2 + (0.1270 - 3.1084)^2 + (0.3811 - 9.3252)^2} = 9.688$$

$$D_{8,1} = \sqrt{(-1.034 - 2.2632)^2 + (0.0384 - 0.3520)^2 + (2.1221 - 3.1084)^2 + (6.3663 - 9.3252)^2} = 4.549$$

Fig. 2. The distance Euclidean values from 1st record of Table1.

Fig.2 shows how to use the KNN model to predict the value of the label column of the first example in the table3. Base on calculating with the KNN model, the distance between the first example in the table3 and the first example

in the table1 must be considered because the calculated result of the $D_{1,1}$ shows that the distance of both examples is nearest.

As the label value of the first example in the label1 is Not Flip, so we also assign the label value for the label column of the first example in table3 as Not Flip too.

$$D_{1,3} = \sqrt{(2.2823 - (-1.0206))^2 + (0.416 - 0.1825)^2 + (2.8879 - 1.2706)^2 + (8.6636 - 3.8117)^2} = 6.093$$

$$D_{2,3} = \sqrt{(0.1822 - (-1.0206))^2 + (0.2811 - 0.1825)^2 + (2.1530 - 1.2706)^2 + (6.4589 - 3.8117)^2} = 3.040$$

$$D_{3,3} = \sqrt{(-1.0650 - (-1.0206))^2 + (0.3858 - 0.1825)^2 + (2.7745 - 1.2706)^2 + (8.3235 - 3.8117)^2} = 4.760$$

$$D_{4,3} = \sqrt{(-0.6414 - (-1.0206))^2 + (0.3454 - 0.1825)^2 + (1.2585 - 1.2706)^2 + (3.7754 - 3.8117)^2} = 1.002$$

$$D_{5,3} = \sqrt{(-0.41762 - (-1.0206))^2 + (0.0204 - 0.1825)^2 + (1.0251 - 1.2706)^2 + (3.0752 - 3.8117)^2} = 0.996$$

$$D_{6,3} = \sqrt{(-0.3048 - (-1.0206))^2 + (0.0234 - 0.1825)^2 + (1.8689 - 1.2706)^2 + (5.6067 - 3.8117)^2} = 2.029$$

$$D_{7,3} = \sqrt{(0.0376 - (-1.0206))^2 + (0.2212 - 0.1825)^2 + (0.1270 - 1.2706)^2 + (0.3811 - 3.8117)^2} = 3.768$$

$$D_{8,3} = \sqrt{(-1.034 - (-1.0206))^2 + (0.0384 - 0.1825)^2 + (2.1221 - 1.2706)^2 + (6.3663 - 3.8117)^2} = 2.697$$

Fig. 3. The distance Euclidean values from 5th record of Table1.

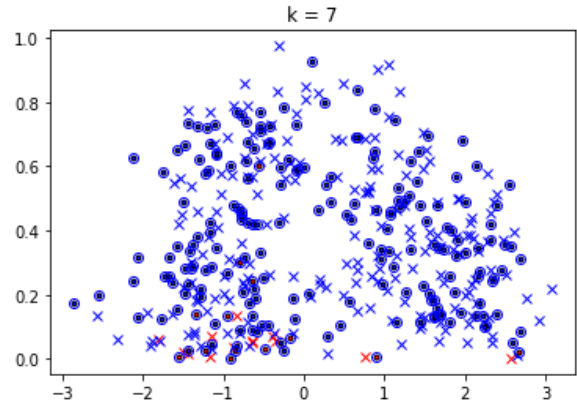
Fig.3 shows how to use the KNN model to predict the value of the label column of the third example in the table3. We use the same method as mentioned earlier. Then we assign the value Flip for the third example in table3 because its data is closest to the fifth example in table1 as shown in the result of $D_{5,3}$.

6. Simulation Result

This work presents the KNN algorithm for creating the model to predict the label value with Flip or Not Flip. The system methods are based on two-dimensional (2D) modulation coding constraints in BPMR system. The target for flip bit conditional can be decreased ITI in the BPMR system, and the performance can be simulated in a future work. This work shows the performance gain of accuracy percentage of models that we compare several systems as Fig.4, Fig.5, and Fig.6.

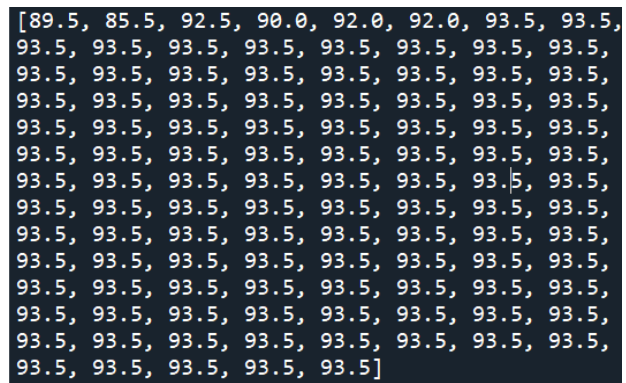
In Fig.4, the system with SNR equal to 2 dB means the system includes high effect noise to decrease the performance, and the accuracy percentage of this system offers the highest accuracy percentage as 93.5% with $k=7$. Where k is the number of values to be compared with the data test.

In Fig.5, the system with SNR equal to 5 dB means the system includes medium effect noise to decrease the performance, and the accuracy percentage of this system offers the highest accuracy percentage in 94.5% with $k=7$.



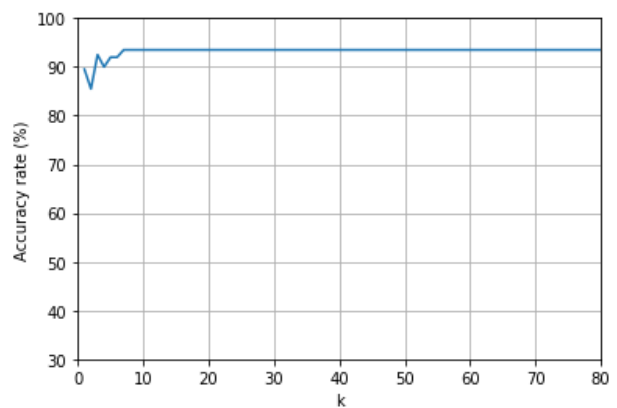
(1a)

The training and testing data of $k=7$ and SNR=2 dB.



(2a)

The value of accuracy (%) of $k=7$ and SNR=2 dB.



(3a)

The value of accuracy (%) of $k=7$ and SNR=2 dB.

Fig 4. The accuracy percentages with SNR =2 dB, K=7 by using K-Nearest Neighbors (KNN) algorithm.

Fig.6 shows the system with SNR equal to 10 dB, which means the system includes low effect noise to decrease the performance, and the accuracy percentage of this system offer 98.0% with k=3. The last result shows the relation between SNR=2 and SNR=5 and SNR=10 dB; it can confirm this algorithm is correct in the accuracy percentage depending on noise in the system, and then, we can generate the BER performance system in future work.

So, comparing accuracy percentage levels that show the high accuracy with each SNR depends on the noise of the SNR value. This conclusion These results have a significant for finding the BER performance in the future to confirm this high accuracy percentage can also be offered the high-performance gain. The Accuracy percentage of the K-Nearest Neighbors Classifier is shown in Table 4.

Table 4. The Accuracy percentage of K-Nearest Neighbors Classifier with each SNR levels.

Accuracy percentage of K-Nearest Neighbors Classifier	
SNR Levels (in dB)	Accuracy (%)
2	93.5
5	94.5
10	98.0

7. Conclusion

The data preparation of this method consists of numerical values. It is appropriate to

use the KNN-algorithm that calculates the numerical values. The KNN-algorithm model can offer the high-performance gain of all SNR levels, and it has some significance to use this algorithm for finding the appropriate of another data set to simulate this model and the full system until getting the BER performance of the BPMR system in the future.

References

- [1] Hai Li, (2016), Storage Physics and Noise Mechanism in Heat-Assisted Magnetic Recording, Doctor of Philosophy, Electrical and Computer Engineering, Carnegie Mellon University, pp. 3-6.
- [2] S. Nabavi, B.V.K.V. Kumar, (2007), Two-dimensional generalized partial response equalizer for bit-patterned media, in: 2007 IEEE International Conference on Communications, pp. 6249–6254.
- [3] A. Arrayangkool, C. Warisarn, P. Kovintavewat, (2013), A recordedbit patterning scheme with accumulated weight decision for bit-patterned media recording, IEICE Trans. Electron. E96-C (12) 1490–1496.
- [4] A. Arrayangkool, C. Warisarn, (2015), A two-dimensional coding design for staggered islands bit-patterned media recording, J. Appl. Phys. 117 (17) 17A904.

- [5] C. Warisarn, A. Arrayangkool, P. Kovintavewat, (2015), An ITI mitigation 5/6 modulation code for bit-patterned media recording, *IEICE Trans. Electron.* E98-C (6) 528–533.
- [6] W. Busyatras, C. Warisarn, S. Koonkarnkhai, and P. Kovintavewat, (2021), A simple 2D modulation code in single-reader two-track reading BPMR systems, *Digital Communications and Networks*, 7 pp. 100-106.
- [7] W Busyatras and C Warisarn, (2020), Soft-information flipping scheme based on a priori LLRs summation for ultra-high density magnetic recording, *AIP Advances*, 10(2), pp. 025217.
- [8] S. Han, G. Kong, and S. Choi, (2019), A Detection Scheme with TMR Estimation Based on Multi-Layer Perceptrons for Bit Patterned Media Recording, *IEEE TRANSACTIONS ON MAGNETICS*, VOL. 55, NO. 7, pp. 3100704.
- [9] S. Jeong and J. Lee, (2022), Soft-Output Detector Using Multi-Layer Perceptron for Bit-Patterned Media Recording, *Applied science*, 12, pp. 2-8.
- [10] S. Jeong and J. Lee, (2022), Bit-Flipping Scheme Using K-Means Algorithm for Bit-Patterned Media Recording, *IEEE Transactions on Magnetics*, DOI 10.1109/TMAG.2022.3144394, pp. IPA-03.
- [11] Seongkwon Jeong, jaejin lee, (2017) “LDPC product coding scheme with extrinsic information for bit patterned media recoding” *AIP Advances* Volume 7, Issue 5, 056513-4.
- [12] C. Warisarn, W. Busyatras, and L. M. M. Myint Soft-(2018) “information flipping approach in multihead multi-track BPMR systems” *AIP Advances* Volume 8, Issue 5, 056509-5.
- [13] N. Rueangnetr, Lin M. Myint, C. Warisarn Soft-(2021)” information flipper based on longshort term memory networks for ultra-high density magnetic recording” *AIP Advances* Volume 11, Issue 1, 015307-6.
- [14] Santi Koonkarnkhai, Chanon Warisarn and Piya Kovintavewat (2021)” A simple skew angle detection and suppression method for bit-patterned magnetic recording” *AIP Advances* 11 Volume 11, Issue 1, 015229-4.