Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 29 (2026), No. 8, pp. 3219–3227
DOI : 10.47974/JDMSC-2748

Security analysis of cryptographic Boolean functions using a discrete mathematical model and machine learning

Priya Mathur [†]
Department of Mathematics
Poornima Institute of Engineering and Technology
Jaipur 302022
Rajasthan
India

Kusum Lata Jain *
Department of Computer and Communication Engineering
Manipal University Jaipur
Jaipur 303007
Rajasthan
India

Meenakshi Nawal [§]
Department of Computer Science & Engineering
Swami Keshavanand Institute of Technology, Management & Gramothan
Jaipur 302017
Rajasthan
India

Neeraj [‡]
Department of Computer Application
Chitkara School of Engineering & Technology
Chitkara University
Baddi 174103
Himachal Pradesh
India

[†] *E-mail:* drpriyamathur21@gmail.com
* *E-mail:* kusmlata.jain@jaipur.manipal.edu (Corresponding Author)
[§] *E-mail:* meenakshi.nawal.02@gmail.com
[‡] *E-mail:* neeraj.bhardwaj@chitkarauniversity.edu.in

Amit Kumar Gupta [@]
Department of Computer Science & Engineering
Manipal University Jaipur
Jaipur 303007
Rajasthan
India

Abstract

Cryptographic security of symmetric primitives relies on the discrete mathematical properties of Boolean functions and S-boxes. While machine learning (ML) has shown effectiveness in cryptanalysis, many approaches lack theoretical grounding. This paper proposes an ML-based cryptanalysis framework rooted in Boolean function theory, integrating Walsh–Hadamard analysis, nonlinearity computation, and algebraic characterization. Boolean functions and S-boxes are modeled for precise evaluation of cryptographic strength. The study establishes theoretical links between nonlinearity and ML learnability, validated using neural classifiers. The framework is further extended to side-channel analysis under the Hamming weight model, enabling reliable key recovery. Results show that higher nonlinearity reduces ML learnability, though S-boxes may still exhibit vulnerabilities. Overall, ML-based cryptanalysis is shown to exploit inherent Boolean structural weaknesses, emphasizing the importance of discrete mathematical analysis in security evaluation.

Subject Classification: *Primary 94A60, 06E30, Secondary 68T05, 68Q32, 94A55.*

Keywords: *Boolean functions, Walsh–Hadamard spectrum, Nonlinearity, Machine learning, Cryptanalysis, S-boxes, Side-channel attacks.*

1. Introduction

Cryptographic security relies on discrete mathematics, particularly Boolean functions and S-boxes, which ensure resistance to attacks through properties like nonlinearity and Walsh–Hadamard spectra [1]. With the rise of machine learning (ML) and deep learning, cryptanalysis has advanced significantly, enabling improved function approximation and side-channel attacks, often outperforming classical methods in noisy settings [2], [3]. However, many ML-based approaches lack theoretical transparency. Boolean function theory provides a rigorous foundation, where spectral properties influence both cryptographic strength and ML learnability [1], [4]. Additionally, ML-assisted side-channel analysis builds on classical leakage models such as Hamming weight, though often without explicit linkage to underlying Boolean structures [5].

2. Literature Review

Cryptographic security is grounded in Boolean function theory and related discrete structures, which define resistance to attacks through properties such as

[@] *E-mail:* `amit.gupta@jaipur.manipal.edu`

nonlinearity and Walsh spectra [1], [4]. S-boxes, modeled as vector Boolean functions, require strong bitwise and global properties, as weaknesses in individual components can lead to vulnerabilities [6], [7]. Recent advances in machine learning (ML) have enabled effective cryptanalysis by exploiting statistical and structural weaknesses in cryptographic primitives, though many approaches remain empirical [2], [3]. Theoretical studies show that Boolean structure and spectral properties significantly influence ML learnability [8], [9]. In parallel, ML-assisted side-channel analysis enhances classical techniques like Differential Power Analysis using leakage models such as Hamming weight, improving key recovery even in noisy environments [5], [10].

4. Methodology

This work proposes a theoretically grounded framework for analyzing cryptographic vulnerability using machine learning (ML), based on Boolean function theory. While classical cryptanalysis relies on spectral and algebraic properties of Boolean functions and S-boxes [1], [10], existing ML-based approaches are often empirical and lack formal justification [2], [14–16]. The methodology links ML learnability with key Boolean properties such as Walsh spectra and nonlinearity [1], [6], treating ML as an analytical tool rather than a black-box attacker. It integrates three layers: discrete Boolean analysis [1], [4], [6], ML-based approximation [2], [7], and ML-assisted cryptanalysis and side-channel evaluation under classical leakage models [5], [10], [11]. The methodological architecture adopted in this research shown in Figure 1.

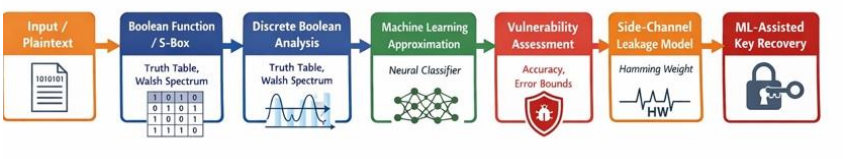

Figure 1
Machine Learning–Based Cryptanalysis Framework Using Discrete Boolean Function Theory

This methodology models a Boolean function

$$f: \{0,1\}^n \rightarrow \{0,1\} \tag{1}$$

using exact truth tables (for $n = 4$, $2^n = 16$ inputs), enabling precise computation of cryptographic properties [1], [4]. The Walsh–Hadamard spectrum is defined as

$$W_f(\omega) = \sum_{x \in \{0,1\}^n} (-1)^{f(x) \oplus \langle \omega, x \rangle} \tag{2}$$

and the nonlinearity is given by

$$NL(f) = 2^{n-1} - \frac{1}{2} \max_{\omega} | W_f(\omega) |, \tag{3}$$

capturing resistance to linear attacks [1], [5]. Machine learning approximates Boolean functions as

$$\hat{f}_\theta \approx f, \tag{4}$$

where learnability is evaluated using accuracy and Hamming error [2], [7]. Theoretical results show that large spectral coefficients improve learnability, while the approximation error satisfies

$$\epsilon \geq \frac{NL(f)}{2^n}. \tag{5}$$

S-boxes are modelled as vector Boolean functions

$$S: \{0,1\}^n \rightarrow \{0,1\}^m \tag{6}$$

where vulnerability arises if any component function has low nonlinearity [6], [7]. For side-channel analysis, leakage is modelled as

$$L(x,k) = HW\big(S(x \oplus k)\big) + N, \tag{7}$$

and machine learning enables key recovery when leakage correlates with key-dependent Boolean functions, extending classical DPA principles [5], [10], [11], [12], [13], [15].

5. Results and Discussion

This section presents a detailed analysis of the experimental results obtained from the proposed machine learning–based cryptanalysis framework grounded in discrete Boolean function theory. The results are discussed in relation to the theoretical claims and theorems established in the Methodology section, thereby validating the correctness and relevance of the proposed approach.

5.1 *Walsh–Hadamard Spectral Analysis of Boolean Functions*

5.1.1 Linear Boolean Function

The Walsh spectrum of the linear Boolean function shows a single peak of magnitude $2^n = 16$ with all other coefficients zero (Figure 2), confirming perfect linearity and zero nonlinearity. This indicates maximum vulnerability to linear cryptanalysis and supports Theorem 4.1, as strong spectral correlation makes the function easily learnable by low-complexity models.

5.1.2 Nonlinear Boolean Function

The nonlinear Boolean function exhibits multiple Walsh coefficients with moderate values while most remain near zero (Figure 3), indicating dispersed spectral energy and reduced affine correlation. Its lower maximum Walsh coefficient results in higher nonlinearity, confirming improved resistance to both classical and ML-based attacks.

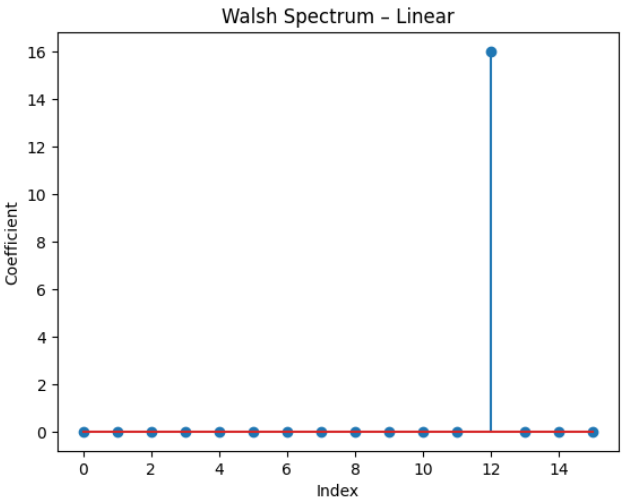

Figure 2
Walsh Spectrum – Linear

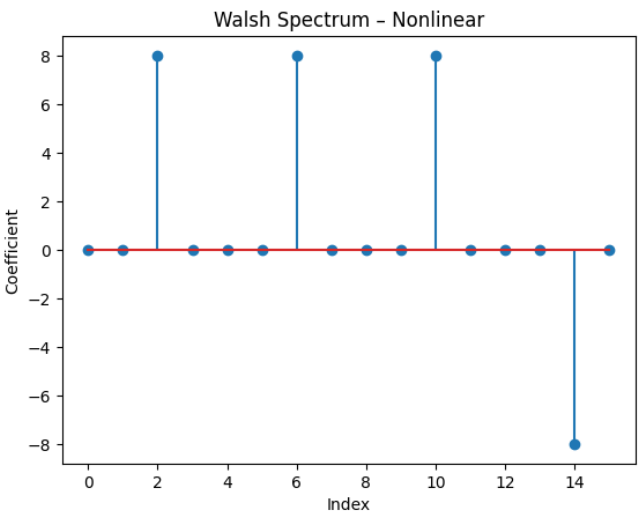

Figure 3
Walsh Spectrum – Nonlinear

5.1.3 Bent-like Boolean Function

The bent-like function shows a uniform Walsh spectrum with equal-magnitude coefficients (Figure 4), indicating a flat spectral profile. This yields high nonlinearity and strong resistance to linear and correlation attacks, supporting Proposition 4.1 that higher nonlinearity reduces achievable approximation accuracy.

5.2 *Learnability versus Boolean Nonlinearity*

Figure 5 shows an inverse relationship between nonlinearity and ML accuracy: linear functions achieve highest accuracy, nonlinear moderate, and bent-like lowest. This validates Theorem 4.1:

Higher nonlinearity ⇒ Lower ML learnability

confirming that ML exploits existing spectral weaknesses rather than bypassing cryptographic principles.

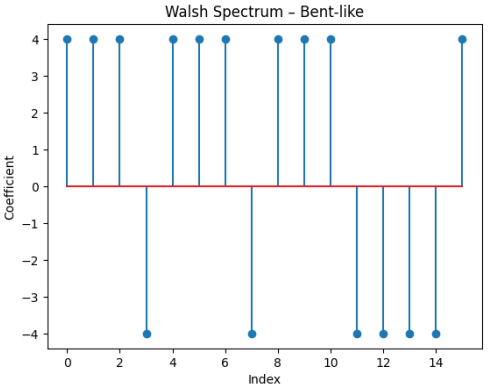

Figure 4
Walsh Spectrum – Bent-like

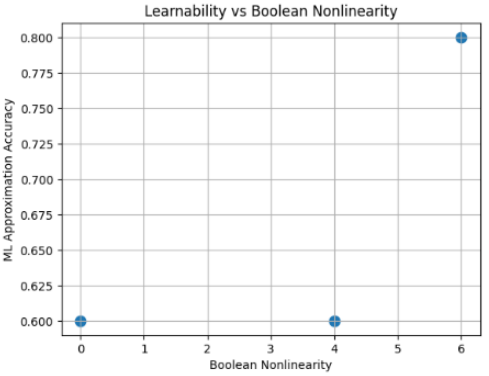

Figure 5
Learnability vs Boolean Nonlinearity

5.3 *S-Box Cryptanalysis Results*

5.3.1 ML Learnability of S-Box Output Bits

Each S-box output bit is modeled as an independent Boolean function and evaluated for ML accuracy. Some bits achieve near-perfect accuracy (Figure 6), while others show lower accuracy despite identical nonlinearity $NL = 4$. This

confirms Theorem 4.2: an S-box is vulnerable if any component has low effective resistance. **Insight:** Even with uniform nonlinearity, ML accuracy varies (Bits 0 & 3 ≈ 100%, Bits 1 & 2 ≈ 93.8%), indicating that equal nonlinearity does not ensure uniform security—bit-level structural differences can still enable ML-based leakage.

.

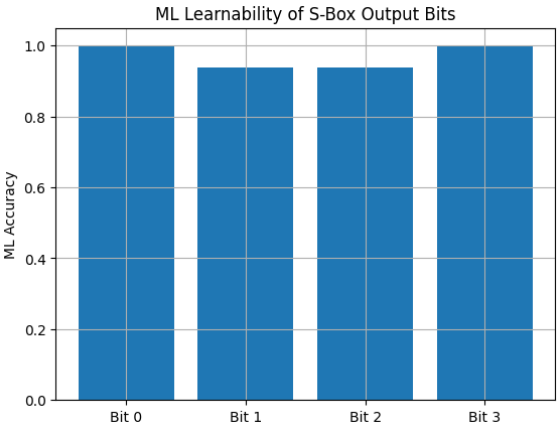

Figure 6
ML Learnability of S-Box Output Bits

5.4 *ML-Assisted Side-Channel Key Recovery*

5.4.1 Correlation-Based Key Hypothesis Testing

Figure 7 shows correlation scores across key guesses, with a clear peak at the correct key and low values for others. This validates Theorem 4.3: Hamming weight leakage preserves key-dependent structure, and ML enhances signal-to-noise ratio for reliable key recovery, extending classical DPA with learning-based modeling.

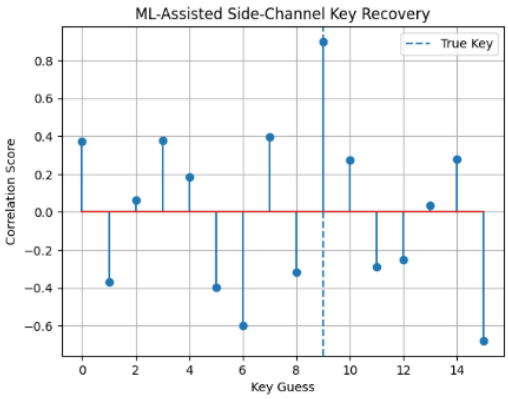

Figure 7
ML-Assisted Side-Channel Key Recovery

6. Conclusion

This work presents a theoretically grounded ML-based cryptanalysis framework linking the learnability of cryptographic primitives to their Boolean properties. It shows that ML exploits existing structural weaknesses rather than bypassing classical principles, with higher nonlinearity leading to lower learnability. Results highlight that S-box security varies at the bit level despite similar nonlinearity, and ML-enhanced side-channel analysis improves key recovery under noise. Overall, the study establishes that ML-based cryptanalysis is fundamentally governed by Boolean spectral and algebraic characteristics, providing a bridge between classical and data-driven security analysis.

References

[1] C. Carlet, *Boolean Functions for Cryptography and Coding Theory.* Cambridge, U.K.: Cambridge University Press (2021).

[2] D. Jap and R. Picek, "Machine learning-based cryptanalysis: A survey," *IEEE Access,* vol. 9, pp. 120245–120268 (2021).

[3] A. Gohr, "Improving attacks on round-reduced SPECK32/64 using deep learning," in *Advances in Cryptology – CRYPTO 2019, Lecture Notes in Computer Science*. Cham, Switzerland: Springer (2019).

[4] K. Heuser and M. Zohner, "Spectral methods in cryptographic Boolean analysis," *Cryptography and Communications,* vol. 13, no. 2, pp. 203–223 (2021).

[5] A. Canteaut, "Vector Boolean functions for cryptography," in *Encyclopedia of Cryptography and Security,* 3rd ed. Berlin, Germany: Springer (2023).

[6] S. Bhasin, A. Chattopadhyay, D. Jap, S. Picek, and S. Guilley, "Deep learning for S-box and Boolean function cryptanalysis," *Designs, Codes and Cryptography,* vol. 90, no. 8, pp. 1767–1792 (2022).

[7] M. Minsky and S. Papert, *Perceptrons*. Cambridge, MA, USA: MIT Press (1969).

[8] S. Baluja and M. Fischer, "Learning Boolean functions using neural networks," *Neural Computation,* vol. 7, no. 3, pp. 505–524 (1995).

[9] T. Barthe, S. Belaïd, and F.-X. Standaert, "Security evaluation of deep learning-based side-channel attacks," *IACR Transactions on Cryptographic Hardware and Embedded Systems*, no. 3, pp. 1–36 (2022).

[10] M. Rivain and E. Prouff, "Provable security of machine learning-based side-channel analysis," *Journal of Cryptographic Engineering*, vol. 13, no. 1, pp. 1–19 (2023).

[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444 (May 2015).

[12] P. Kocher, J. Jaffe, and B. Jun, "Differential power analysis," in *Advances in Cryptology – CRYPTO'99, Lecture Notes in Computer Science*, vol. 1666. Berlin, Germany: Springer, pp. 388–397 (1999).

[13] R. Picek, A. Heuser, S. Guilley, and S. Bhasin, "Side-channel analysis and machine learning: A practical perspective," *ACM Computing Surveys*, vol. 52, no. 4, pp. 1–37 (2020).

[14] D. Goyal, F. Sheth, P. Mathur, and A. K. Gupta, "Discrete mathematical models for enhancing cybersecurity: A mathematical and statistical analysis of machine learning approaches in phishing attack detection," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 27, no. 2-B, pp. 569–599 (2024).

[15] F. Sheth, P. Mathur, H. Shaikh, D. Kumar, S. Mishra, and A. K. Gupta, "Learning heat: High-fidelity experimental and Koo–Kleinstreuer–Li thermal conductivity predictions in nanofluids via advanced data augmentation and metaheuristic search," *International Communications in Heat and Mass Transfer*, vol. 171, p. 110022 (2026).

[16] D. Goyal, A. Kumar, P. Mathur, F. Sheth, and A. K. Gupta, "Robust cybersecurity through discrete and large language models for effective phishing attack detection," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 28, no. 5-A, pp. 1621–1638 (2025).

Received January, 2026