Research Article

International Journal of Data Science and Applications (JOINDATA) 8(1), 1-10, 2025

Received: 4-Oct-2024 Accepted: 28-Dec-2024

homepage: https://dergipark.org.tr/tr/pub/joindata

SAKARYA UYGULAMALI BİLİMLER ÜNİVERSİTESİ

Neural Network-Based Approaches to High-Energy Physics

Serpil YALÇIN KUZU[1*], Ayben KARASU UYSAL[2], Mustafa KAYA[3]

[1] Department of Physics, Faculty of Science, Firat University, skuzu@firat.edu.tr

[2,] Department of Physics, Faculty of Science, Yildiz Technical University, ayben.karasu@yildiz.edu.tr

[3] Department of Physics, Faculty of Science, Firat University, mustafakaya23100@gmail.com

ABSTRACT

The exploration of quarkonium states at the Large Hadron Collider (LHC) plays a critical role in advancing particle physics and validating quantum chromodynamics (QCD). This study focuses on the decay of J/ψ mesons into electron-positron pairs ($J/\psi \rightarrow e^+e^-$), a process that offers valuable insights into the dynamics of high-energy collisions but poses challenges due to the vast and complex datasets generated. To address these challenges, Deep Neural Networks (DNNs) was applied employing techniques such as data preprocessing, feature engineering, and hyperparameter tuning to improve the efficiency and accuracy of event classification. The analysis utilized a dataset from the CMS experiment at $\sqrt{s}$ = 7 TeV, focusing on features such as transverse momentum, azimuthal angle differences, and particle opening angles. The DNN model demonstrated exceptional performance, achieving a precision of 97.5%, a recall of 99.1%, and an F1-score of 98.3%. The area under the receiver operating characteristic curve (AUC ROC) was 99.8%, underscoring the model's robustness in distinguishing signal events from background noise. These results highlight the potential of DNNs in advancing particle identification and interpreting high-energy physics data. The study concludes that DNNs are a powerful tool for quarkonium state analysis and can significantly enhance the precision of particle measurements, paving the way for broader applications in high-energy physics research.

Keywords: Quarkonium, charmonium, deep neural networks, high energy physics.

Yüksek Enerji Fiziğinde Yapay Sinir Ağı Tabanlı Yaklaşımlar

ÖZ

Büyük Hadron Çarpıştırıcısı (Large Hadron Collider - LHC)’nda kuarkonyum durumlarının incelenmesi, parçacık fiziğinin ilerlemesi ve kuantum renk dinamiği (QCD) kuramının doğrulanması açısından kritik bir rol oynamaktadır. Bu çalışma, J/ψ mezonlarının elektron-pozitron çiftlerine bozunumu ($J/\psi \rightarrow e^+e^-$) üzerine odaklanmaktadır. Bu bozunma süreci, yüksek enerjili çarpışmaların dinamiklerine dair değerli bilgiler sunarken, üretilen büyük ve karmaşık veri kümeleri nedeniyle analizleri oldukça karmaşıktır. Bu zorlukların üstesinden gelmek amacıyla, olay

[*] Corresponding Author’s email: skuzu@firat.edu.tr

sınıflandırma verimliliğini ve doğruluğunu artırmak için veri ön işleme, öznitelik mühendisliği ve hiperparametre ayarlaması gibi tekniklerin uygulandığı Derin Sinir Ağları (DNN) kullanılmıştır. Analizde, $\sqrt{s} = 7$ TeV enerjisinde çalışan Kompakt Müon Selenoid (Compact Muıon Solenoid – CMS) deneyi veri kümesi kullanılmış ve analizde özellikle enine momentum, açısal farklar ve parçacıklar arasındaki açıklık açısı gibi özniteliklere odaklanılmıştır. Geliştirilen DNN modeli %97.5 doğruluk, %99.1 duyarlılık ve %98.3 F1-skoru ile yüksek bir performans sergilemiştir. Ayrıca, modelin sinyal olaylarını arka plan gürültüsünden ayırmadaki başarısını ortaya koyan alıcı işletim karakteristik eğrisi altında kalan alan (AUC ROC) %99.8 olarak hesaplanmıştır. Elde edilen sonuçlar, DNN'lerin parçacık tanımlama süreçlerinde ve yüksek enerjili fizik verilerinin yorumlanmasında büyük potansiyele sahip olduğunu göstermektedir. Çalışma, DNN'lerin kuarkonyum durumu analizinde güçlü bir araç olduğunu ve parçacık ölçümlerinin hassasiyetini önemli ölçüde artırarak yüksek enerjili fizik araştırmalarında daha geniş uygulama alanları sunduğunu ortaya koymaktadır.

Anahtar Kelimeler: Kuarkonyum, çarmonyum, derin sinir ağları, yüksek enerji fiziği.

1 Introduction

The integration of machine learning (ML) technologies has revolutionized various scientific disciplines, including particle and high-energy physics. Among these, the Large Hadron Collider (LHC), the world's largest and most powerful particle accelerator, stands out for generating unprecedented volumes of data, creating both opportunities and challenges for researchers [1, 2]. Traditional analysis techniques, while effective for smaller datasets, struggle to cope with the complexity and scale of modern high-energy physics data [3]. These limitations have led to a growing reliance on advanced methodologies, such as neural networks, to unlock new insights.

Recent advancements in deep learning, a subset of machine learning, have proven particularly transformative. Deep Neural Networks (DNNs) can automatically learn representations from raw data, identify intricate patterns, and make predictions with high accuracy [4]. Their ability to manage large, complex datasets makes them a powerful tool for analysing quarkonium states, such as J/ψ mesons. These particles play a pivotal role in validating quantum chromodynamics (QCD), the theory describing strong interactions between quarks and gluons [5, 6]. The decay of J/ψ mesons into electron-positron pairs ($J/\psi \rightarrow e^+e^-$) offers a unique opportunity to study high-energy collision dynamics, but analyzing these decays requires robust methods to sift through extensive and noisy datasets [2, 5].

While traditional analysis methods depend heavily on manual feature selection, introducing potential systematic uncertainties, DNNs provide a data-driven alternative. They minimize reliance on manual cuts and reduce uncertainties through advanced preprocessing and feature engineering techniques [3, 7]. The ability of DNNs to effectively distinguish signal from background events has inspired a broad spectrum of applications in high-energy physics, from particle identification to collision event classification [8, 9]. Recently, researchers have demonstrated novel ML-based frameworks tailored for High-Luminosity LHC data processing, significantly improving efficiency and scalability [10, 11].

Studies from the last two years have highlighted the transformative potential of quantum-inspired machine learning in high-energy physics. Methods leveraging quantum data representation and machine learning models such as quantum convolutional neural networks have shown the capacity to improve classification accuracy and model interpretability in high-dimensional datasets [10, 11]. Furthermore, advancements in symbolic computation frameworks and the use of transformers for squared amplitude computations offer unprecedented speed and accuracy, paving the way for broader applications in physics [7, 12].

This study focuses on the application of DNNs to the analysis of charmonium states, specifically J/ψ mesons. By employing techniques such as data preprocessing, hyperparameter tuning, and advanced evaluation metrics, this research highlights the efficacy of DNNs in particle physics. The results demonstrate significant advancements in signal-to-noise discrimination and offer a roadmap for future applications in high-energy physics research.

2 Research Methodology

2.1 Deep Neural Networks (DNNs)

Deep Neural Networks (DNNs) consist of multiple layers of neurons, where each layer transforms input data into more abstract and useful representations, making them well-suited for tasks that involve large and noisy datasets, such as those produced by LHC experiments [1-3]. The application of DNNs involves several key steps, starting with data preprocessing and feature selection. The architecture of the DNNs typically includes an input layer, multiple hidden layers with nonlinear activation functions like The Rectified Linear Unit (ReLU) activation function [13], and an output layer that uses a Sigmoid function [14] for classification tasks. This layered structure allows the DNNs to capture complex patterns in the data, which is crucial for identifying specific events, such as the decay of J/ψ mesons [3, 4]. The training process involves compiling the DNNs with a binary cross entropy function, which is suitable for binary classification tasks. In the algorithm an optimizer like Adam [15] is used for its efficiency in handling sparse gradients to prevent overfitting.

Algorithm 1 DNNs for 6 layers including 1 input, 4 hidden, and 1 output layers.

Input: Preprocessed LHC data (X), Labels (y)

Output: Trained DNNs Model (DNNs)

Step 1: Initialize DNNs with architecture:

Input Layer: size = number of features in X

Hidden Layers:

Layer 1: 512 neurons, ReLU activation

Layer 2: 256 neurons, ReLU activation

Layer 3: 128 neurons, ReLU activation

Layer 4: 64 neurons, ReLU activation

Output Layer: size = number of classes in y, Sigmoid activation

Step 2: Compile DNNs with:

Loss function = Binary Cross-Entropy

Optimizer = Adam

Metrics = [Precision, Recall, F1-score, AUC ROC]

Step 3: Split X and y into training set (X_train, y_train)

Step 4: Train DNNs on X_train, y_train with:

Batch size = 128

Epochs = 100

Step 5: Evaluate DNNs on test set if available

Return DNNs

The pseudo-algorithm for the DNNs process can be summarized in Algorithm 1. In this pseudo-algorithm, important parameters such as the number of neurons in each hidden layer, the learning rate, and the batch size are carefully chosen to optimize the model's performance. The ReLU activation function is employed in the hidden layers due to its efficiency and ability to introduce non-linearity, which is crucial for learning complex patterns. The Adam optimizer, known for its adaptability and efficiency, is used to minimize the loss function during training. Selection of epoch number and batch size are key parameters to improve the model performance. To prevent overfitting, the Dropout method can be used in the algorithm during training that randomly removes a fraction of neurons, which forces the network to develop more robust features by improving generalization ability of the DNNs [16].

2.2 Experimental Design

2.2.1.1 Dataset

In this study proton–proton collision data collected by the CMS experiment at $\sqrt{s} = 7$ TeV in 2010 were used to reconstruct *J/ψ* from its dielectron decay channel [17-19]. The corresponding channel was investigated from 2 to 5 GeV.

2.2.1.2 Preprocessing and Feature Selection

In the analysis of J/ψ meson identification, careful preprocessing and feature selection are critical to enhance the accuracy of the DNNs model. For this purpose, the data having missing information were removed from the dataset. 10015 events were analyzed with Python [20]. For DNNs application Keras [21] and TensorFlow [22] packages were used. Extracting the particle signal distribution from the extensive background is a significant challenge in particle physics analysis due to the overlap between these distributions. Traditional analysis methods typically involve applying specific selection cuts on particle features to reject most of the background. However, these cuts often lead to the loss of particle signal information and introduce high systematic uncertainties. In contrast, the application of the DNNs leverages the inherent power of each feature to discriminate the signal, thus minimizing the need for such cuts and reducing the associated uncertainties.

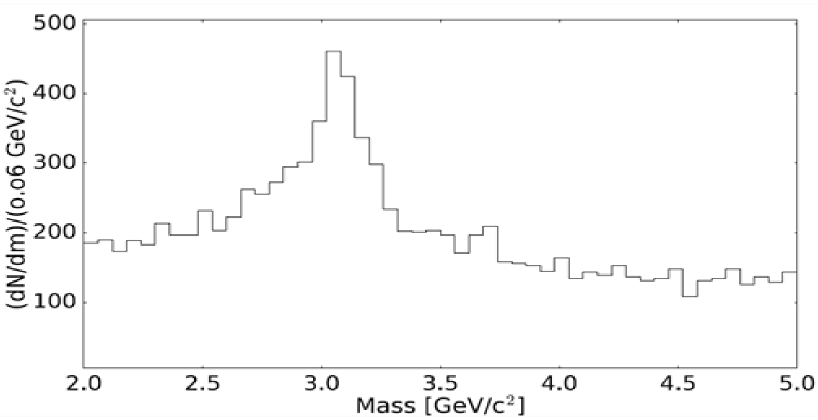

Figure 1. Mass distribution of pairs in the dataset

In the analysis, information of each decay partner, including momentum components (p_x, p_y, p_z), transverse momentum (pt), energy (E), pseudorapidity (η), azimuthal angle (ϕ), the charges (q), and the mass (M) of the pairs were provided by the dataset. Mass distribution of the pairs in the dataset is represented in Figure 1. In the analysis instead of using η and ϕ information, the absolute differences between two particles ($\Delta\eta$ and $\Delta\phi$) were calculated and used as features to determine the correlation between particles within the same event [23-25]. Also, the opening angle (α) between the partners [23] and the transverse momentum of the mother particle (P_t) were calculated and used as features.In the study, *J/ψ*, the signal class, defined by pairs including electrons and positrons having track properties such as each partner must have p_t greater than 1.3 GeV/c and charge in opposite sign, which are the regular cuts to increase the performance of the particle measurement [23]. An optimized mass range description is necessary to distinguish signal and background classes. Mass of the meson is 3096.9±0.006 MeV/c^2 [26], therefore, the signal must be on a narrow mass window within ±300 MeV/c^2 of this value in order to have background contamination as low as possible. With these definitions signal and background instances are listed in Table 1.

Table 1. List of signal and background class numbers.

Class Labels	Number of Instances
Signal (*J/ψ*)	3345
Background	8251

2.2.1.3 DNNs Implementation

In this study, 75% of the data was allocated to the training set, while the remaining 25% was reserved for testing. The decision to use a 75/25 split was based on a balance between providing sufficient data for training the DNNs and maintaining an adequate amount for robust testing and evaluation. This proportion is widely used in machine learning studies because it ensures the model is trained on a diverse dataset while still allowing for a meaningful validation of its performance on unseen data [1, 2]. If the training proportion were increased to 80%, it would leave only 20% of the data for testing. While this could potentially improve the training accuracy by exposing the model to more data during training, it might also reduce the reliability of the evaluation metrics due to the smaller test set. Conversely, reducing the training set size below 75% could limit the model's ability to generalize effectively, as it would have fewer samples to learn from.

The features that was used in the DNNs algorithm to identify signal class were selected according to the traditional charmonium analysis. Therefore, α, $\Delta\eta$, $\Delta\phi$, M, P_t, and q and p_t of each pair partner were used in the DNNs to enhance reliability of the model in complex particle dataset.In the DNNs, the Adam algorithm was preferred due to its performance in minimizing the difference between the output value produced by the network and the actual value, as well as its ability to prevent overfitting [15, 27]. The ReLU function was selected as the activation function for the input and hidden layers. The Sigmoid function was used as the activation function for the output layer. Since there are two classes in the dataset, 'binary cross entropy' was preferred as the loss function. To prevent overfitting in the model, the layers were organized using the Dropout method with a dropout rate of 0.2, and the kernel initializer [28] suitable for the network structure was set to 'normal'.

The specific configuration of the network was determined through an iterative process involving trial-and-error and hyperparameter tuning. This process aimed to maximize the model's performance metrics (e.g., precision, recall, F1-score, and AUC ROC) while minimizing overfitting and ensuring computational efficiency. Alternative configurations were tested, and the chosen architecture demonstrated the best balance between accuracy and generalization ability. In the end of the optimization process, the DNNs was structured with a total of 5 layers: one input layer, one output layer, and three hidden layers. The number of nodes for each layer was set to 9, 150, 150, 20, and 1, respectively. Based on the studies, the batch size, which is the number of samples to be fed into the network at once during the learning process, was determined 128, and the number of epochs to be 50.

3 Results and Discussion

3.1 ML Metrics

The performance of the DNNs were evaluated by using ML assessment metrics. The metrics are precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC ROC), that are used to demonstrate the capability of the model to determine the particles in complex high energy physics datasets. Precision, recall, and F1 score values can be determined using the parameters found in the confusion matrix [29] obtained from the application of the DNNs model. The confusion matrix is a table that shows the actual class information in the dataset alongside the predictions made by the DNNs model. An example of confusion matrix was is represented in Table 2.

Table 2. An example of confusion matrix.

		Predicted	
	Classes	*Negative*	*Positive*
Actual	*Negative*	True Negative (TN)	False Positive (FP)
	Positive	False Negative (FN)	True Positive (TP)

The actual values represent the class information from the dataset, while the predicted values are the class information obtained as a result of the model's detection. In the confusion matrix used in the study, the positive class represented the J/ψ meson signal class, and the negative class was the background class. The confusion matrix was used to determine how the applied model categorized the data. From the confusion matrix true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values were extracted for calculation of the metrics. In the analysis true positive (TP) indicates the data belonging to the correctly identified J/ψ meson signal by the DNNs; true negative (TN) indicates the data belonging to the correctly determined background class by the model; false negative (FN) indicates the incorrectly labeled J/ψ meson signal by the DNNs; and false positive (FP) indicates the data belonging to the incorrectly labelled background by the model. In an ideal scenario, it is expected that the FP and FN values is 0. Recall or sensitivity is the ratio of the correctly identified J/ψ meson by the model to the total correct signal class in the dataset calculated using the parameters in the confusion matrix as demonstrated in Equation 1 [30]. Precision is the performance of the DNNs in identifying the signal computed by Equation 2 [30]. The F1 score, which allows for an evaluation of the model's performance by considering both sensitivity and precision together, is calculated using the harmonic mean as in Equation 3 [30].

$$Recall = \frac{TP}{TP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{3}$$

The performance of the DNNs was measured using the "Receiver Operating Characteristic (ROC)" method, which combines several metrics. The ROC curve is based on (i) determining the two-dimensional graph of (ii) sensitivity or True Positive Rate (TPR) and (iii) the "false positive rate (FPR = FP / (FP + TN))" metrics and (iv) calculating the area under the curve (AUC ROC). The AUC ROC is a measure of how well the model can distinguish between the two classes [30]. A high AUC ROC value indicates the model's strong ability to differentiate between classes.

3.2 Results

The performance of the DNNS in detecting the dielectron decay channel was evaluated by ML metrics. The confusion matrix of the analysis was demonstrated in Table 3. From the matrix it is revelaed that the model identified 615 J/ψ signal out of 616 correctly from dielectron decay channel. The derived ML metrics are listed in Table 4.

Table 3. The confusion matrix of the DNNs.

		Predicted	
	Classes	*Negative*	*Positive*
Actual	*Negative*	1856 (TN)	32 (FP)
	Positive	1 (FN)	615 (TP)

The ROC of the model is represented in Figure 2. From the table and the AUC ROC value it is concluded that the DNNs algorithm 99.771% successfully determined the meson with high precision and sensitivity. The results reveal the power of the DNNs in particle identification.

Table 4. *The performance of the DNNs.*

Precision	Recall	F1 Score
0.97500	0.99071	0.98253

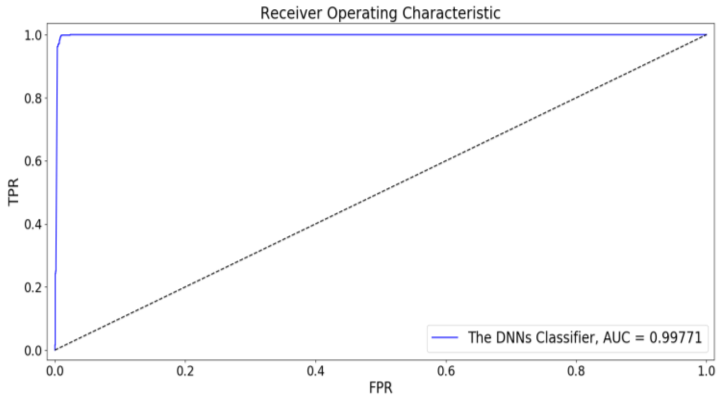

Figure 2. The ROC plot of the DNNs.

4 Conclusions

The integration of advanced machine learning (ML) techniques into high-energy physics significantly enhances the interpretation and analysis of complex datasets. This study demonstrates the successful application of DNNs to identify and analyse the ground state of charmonium, specifically the J/ψ meson, from the extensive datasets generated by the Large Hadron Collider (LHC). The findings confirm that DNNs offer a robust and efficient approach for meson analysis, with the potential to improve particle identification accuracy while reducing systematic uncertainties.

The results underscore the capability of DNNs to handle high-dimensional, noisy data, making them a valuable tool for exploring other particles and phenomena in high-energy physics. This work not only validates the applicability of DNNs in quarkonium state analysis but also suggests broader utility in similar tasks involving large-scale data. By bridging the gap between machine learning and particle physics, this study highlights the transformative impact of ML technologies in advancing the field and for future research aimed at further optimizing these techniques for more complex particle datasets.

Declarations

4.1 Study Limitations

None.

4.2 Acknowledgements

The present study is part of Mustafa Kaya's thesis work towards the fulfillment of his Master's degree at Firat University. A part of this study was performed as an oral presentation at the 6th International Conference on Data Science and Applications (ICONDATA'24). This work was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) Project Number 123F060 and Yildiz Technical University Project Number FBA-2024-6089.

4.3 Competing Interests

There is no conflict of interest in this study.

4.4 Authors' Contributions

Author 1 contributed to the conceptualization and design of the study, performed the data analysis, and drafted the initial manuscript. Author 2 assisted in the study design, contributed to the data interpretation, and provided critical revisions to the manuscript. Author 3 was responsible for the data analysis, and improvement of manuscript. All authors reviewed, revised, and approved the final manuscript for submission.

Serpil YALÇIN KUZU: The author was responsible for the conceptualization and design of the study, performed the data analysis, and drafted the initial manuscript.

Ayben KARASU UYSAL: The author assisted in the study design, contributed to the data interpretation, and provided critical revisions to the manuscript.

Mustafa KAYA: The author assisted the data analysis, and improvement of manuscript.

References

[1] A. Radovic *et al.*, "Machine learning at the energy and intensity frontiers of particle physics," *Nature*, vol. 560, no. 7716, pp. 41–48, 2018.

[2] S. Yalcin Kuzu, "J/ψ production with machine learning at the LHC," *Eur. Phys. J. Plus*, vol. 137, p. 392, 2022.

[3] J. Smith *et al.*, "Application of neural networks in high energy physics: An overview," *Phys. Lett. B*, vol. 824, p. 136834, 2021.

[4] P. Kumar and S. Banerjee, "Deep learning approaches for quarkonium state analysis in high-energy collisions," *Nucl. Phys. B*, vol. 976, p. 115715, 2023.

[5] Y. Zhang and W. Chen, "Enhancing particle identification with deep learning in LHC data," *Comput. Phys. Commun.*, vol. 251, p. 107099, 2020.

[6] M. Trenti *et al.*, "Quantum-inspired machine learning on high-energy physics data," *npj Quantum Inf.*, vol. 7, p. 8, 2020.

[7] L. Giommi *et al.*, "Cloud native approach for Machine Learning as a Service for High Energy Physics," in *Proc. ISGC 2022*, 2022.

[8] A. Butter *et al.*, "Machine learning and LHC event generation," *SciPost Phys.*, vol. 14, no. 4, p. 79, 2022.

[9] Q. Zhao *et al.*, "Δ2 machine learning for reaction property prediction," *Chem. Sci.*, vol. 14, pp. 13392–13401, 2023.

[10] L. Nagano *et al.*, "Quantum data learning for quantum simulations in high-energy physics," *Phys. Rev. Res.*, vol. 5, p. 043250, 2023.

[11] A. Alnuqaydan *et al.*, "SYMBA: Symbolic computation of squared amplitudes in high energy physics with machine learning," *Mach. Learn.: Sci. Technol.*, vol. 4, 2022.

[12] A. Delgado and K. E. Hamilton, "Unsupervised quantum circuit learning in high energy physics," *Phys. Rev. D*, vol. 106, p. 096006, 2022.

[13] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, 2010, pp. 807–814.

[14] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint*, arXiv:1412.6980, 2014.

[16] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[17] T. McCauley, "CMS releases open data for Machine Learning," *CERN News*, 2014. [Online]. Available: https://cms.cern/news/cms-releases-open-data-machine-learning

[18] T. McCauley, "J/ψ to two electrons from 2010," *CERN Open Data Portal*, 2014. [Online]. Available: https://opendata.cern.ch/record/302. DOI: 10.7483/OPENDATA.CMS.97YF.C4AH

[19] T. McCauley, "Events with two electrons from 2010," *CERN Open Data Portal*, 2014. [Online]. Available: https://opendata.cern.ch/record/304. DOI: 10.7483/OPENDATA.CMS.PCSW.AHVG

[20] Python Software Foundation, "Python documentation," [Online]. Available: https://docs.python.org/3/. [Accessed: May 11, 2025].

[21] F. Chollet, "Keras: The Python deep learning library," [Online]. Available: https://keras.io. [Accessed: May 11, 2025].

[22] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Operating Systems Design and Implementation (OSDI)*, 2016, pp. 265–283.

[23] S. Nourbakhsh, "Studio degli eventi J/ψ in due elettroni con i primi dati di CMS," 2010.

[24] NA61/SHINE Collaboration, "Recent results," *Eur. Phys. J. C*, vol. 80, p. 1151, 2020.

[25] CMS Collaboration, "Centrality dependence of high-pT charged particle suppression in PbPb collisions at √sNN = 5.02 TeV," *Phys. Rev. C*, vol. 96, p. 014915, 2017.

[26] M. Tanabashi *et al.*, "Review of Particle Physics," *Phys. Rev. D*, vol. 98, p. 030001, 2018.

[27] N. S. Keskar and R. Socher, "Improving generalization performance by switching from Adam to SGD," *arXiv preprint*, arXiv:1712.07628, 2017.

[28] Y. LeCun *et al.*, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015. DOI: 10.1038/nature14539.

[29] A. C. Müller and S. Guido, *Introduction to Machine Learning with Python*, Sebastopol, CA, USA: O'Reilly Media Inc., 2016.

[30] M. Sokolova *et al.*, "Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation," in *Australasian Joint Conf. Artif. Intell.*, Hobart, Australia, 2006, pp. 1015–1021. DOI: 10.1007/11941439_114.