Bozok Journal of Science
Volume 4, No 1, Page 51-66 (2026)
Research Article DOI: 10.70500/bjs.1944276

Predicting hadron mass and decay width using XGBoost: A unified machine learning framework for baryons and mesons

Metin YALVAÇ[1*], Tarık AKAN[1]

[1]Yozgat Bozok University, Faculty of Science and Letters, Department of Physics, 66100, Yozgat, Türkiye
[*]Correspondence: metin.yalvac@bozok.edu.tr

Received: 05.05.2026 Accepted: 20.05.2026 Final Version: 20.05.2026

Abstract

This study presents an XGBoost-based machine learning framework for predicting the mass and decay width of hadrons, including both baryons and mesons, within a unified feature space. The dataset was compiled from Particle Data Group (PDG) records and contains 495 hadrons in total, reduced to 406 and 373 usable samples for mass and width prediction tasks respectively after data cleaning. Model inputs consist of quark and antiquark counts along with conserved quantum numbers including isospin, spin, parity, electric charge, strangeness, charmness, and bottomness. Analyses were conducted at three levels: baryon-only, meson-only, and combined hadron datasets. For width prediction, two scenarios were evaluated: one excluding and one including hadron mass as an input feature. Mass prediction achieved high accuracy across all datasets, with test R^2 values of 0.904, 0.959, and 0.964 for baryons, mesons, and the combined hadron dataset, respectively. Feature importance analysis identified heavy-flavor quark content, particularly bottom quarks, as the dominant factor in mass prediction. Width prediction without mass information yielded limited explanatory power, especially for mesons (R^2=0.203), whereas including mass as an input substantially improved performance, raising the combined hadron model's test R^2 to 0.812. These results confirm that hadron mass plays a central role in governing decay width through kinematic decay channels. Data augmentation via Gaussian-noise perturbation did not improve generalization over the baseline model trained on real observations alone. The study demonstrates that XGBoost offers both predictive accuracy and physical interpretability, providing a complementary data-driven tool for hadron spectroscopy.

Keywords: Hadron spectroscopy, XGBoost, mass prediction, decay width, machine learning

1. INTRODUCTION

Hadron spectroscopy is one of the most fundamental research areas aimed at understanding the structure of strong interactions in the low-energy regime. Baryons and mesons, as composite systems of quarks, are classified through their quantum numbers, internal quark content, masses, and decay widths (Malekhosseini et al., 2024). These properties play a decisive role both in testing theoretical models and in interpreting the physical nature of newly observed particle candidates. Current compilations by the PDG systematically present experimentally measured features such as mass and width for a vast family of particles, including mesons and baryons, demonstrating that this field provides a robust foundation for data-driven studies (Navas et al., 2024; Workman et al., 2022).

The explanation of hadronic masses and widths has traditionally been addressed within the frameworks of quark models, potential models, lattice QCD calculations, and various phenomenological approaches. However, the highly nonlinear and multivariate relationship between hadron properties and their quark content has made the application of data-driven methods increasingly attractive (Malekhosseini et al., 2024). Particularly in recent years, machine learning has transitioned from simple event classification to the direct prediction of spectral quantities (Gal et al., 2022). Recent studies, such as those by Akan and Malekhosseini et al., have demonstrated that predicting mass and width based on quantum numbers is possible with high accuracy, even for states that are experimentally uncertain or exotic (Akan, 2024; Malekhosseini et al., 2024).

Despite these advancements, a significant portion of the current literature remains meson-centric, and the joint investigation of baryons and mesons under a single "hadron" framework within a common feature space appears relatively limited (Rostami et al., 2025; Tong et al., 2025). Aiming to fill this gap, this study proposes an XGBoost-based prediction framework utilizing a three-level design. By first

[1] https://orcid.org/0000-0003-4915-9162 [2] https://orcid.org/0000-0002-9751-3405

analyzing baryons and mesons separately, the study evaluates whether more homogeneous, subclass-specific patterns can be learned with higher stability and interpretability. Subsequently, by merging both groups into a unified hadron dataset, it tests the extent to which a single model can capture shared regularities across the full hadron spectrum. In this way, the study is positioned not only as a hadron-level prediction task but also as a comparative framework between subclass-based and unified modeling.

Among machine learning techniques, XGBoost emerges as a powerful method that combines the tree-based gradient boosting approach with advantages in scalability, regularization, and sensitivity to sparse data (Chen & Guestrin, 2016). Developed by Chen and Guestrin, this algorithm is particularly advantageous for studies seeking to examine the relative effects of physical variables on target quantities, as it allows for the interpretation of feature importance in addition to producing high accuracy (Chen & Guestrin, 2016; Nielsen, 2016). Considering the complex interactions between quark contents, conserved quantum numbers, and derived features, XGBoost is evaluated as a strong candidate model for predicting quantities such as hadron mass and width.

Input variables such as quark contents, isospin, spin, parity, and electric charge were used. A direct regression model was established for mass prediction, while for width prediction, two different scenarios, one including mass information and one excluding it, were evaluated. Thus, the contribution of hadron mass to width prediction was quantitatively tested, and model behavior was analyzed in detail through correlation analysis, feature importance, residual distributions, and true-vs-predicted value comparisons.

The primary contribution of this study is the development of a hadron-level XGBoost model that treats baryons and mesons together, extending the data-driven approach previously proposed for mesons to a broader physical scope. Such an approach is significant not only for uncovering hidden patterns within existing experimental data but also for providing a supplementary tool that generates predictions for hadronic states with limited information. The main contributions of this study are: development of a unified XGBoost framework treating baryons and mesons jointly for both mass and width prediction, a quantitative demonstration of the role of hadron mass as a dominant input for width prediction, and an interpretable feature importance analysis identifying which quark contents and quantum numbers drive predictive performance across hadron families.

2. MATERIAL AND METHODS

2.1 Dataset

The dataset for this study was constructed as a structured tabular framework, integrating the fundamental quark compositions of hadrons with their experimentally reported physical properties. Compiled primarily from the Particle Data Group (PDG, 2020) summary tables, the compilation provides a systematic mapping of quark content, conserved quantum numbers, masses, and decay widths. The raw data initially encompassed 495 hadrons, 292 baryons and 203 mesons. Examples from dataset are shown in Table 1 & Table 2.

Table 1. Baryon dataset examples used as inputs taken from PDG.

Name	Quark content	J	I	P	Electric charge	Mass (MeV/c^2)	Width (MeV)
p (proton)	uud	1/2	1/2	+	+1	938.272	Stable
n (neutron)	udd	1/2	1/2	+	0	939.565	Stable*
Ξ^0	uss	1/2	1/2	+	0	1314.86	–
Ω_c^0	ssc	1/2	0	+	0	2695.2	–

Table 2. Meson dataset examples used as inputs taken from PDG.

Name	Quark content	J	I	P	Electric charge	Mass (MeV/c^2)	Width (MeV)
π^0	$(u\bar{u} - d\bar{d})/\sqrt{2}$	0	1.0	−	0	135.0	0.0
K^+	$u\bar{s}$	0	0.5	−	+1	494.0	0.0
$\phi(1020)$	$s\bar{s}$	1	0.0	−	0	1019.2305	4.263
$J/\psi(1S)$	$c\bar{c}$	1	0.0	−	0	3096.0	0.093

To ensure statistical consistency across different regression tasks, a task-specific data cleaning strategy was implemented rather than a single global filter. This approach was chosen to maximize the preservation of physically meaningful observations; for instance, a record

missing a width value could still provide valuable information for mass prediction. Consequently, the usable sample size was refined to 406 for mass prediction (205 baryons, 201 mesons) and 373 for width prediction (172 baryons, 201 mesons). This distribution highlights a relative scarcity in baryon width data, whereas the meson subset demonstrates higher representative power for width analysis.

Model input variables were carefully selected to represent internal structures and quantum states. Beyond direct quark and antiquark counts, the feature space includes essential physical quantum numbers such as; Strangeness (S), Charm (C), Bottomness (B'), Electric Charge, Isospin (I), Spin (J), and Parity (P). In unified models, a binary variable was introduced to distinguish between baryon and meson formations. To further explore the relationship between spectral quantities, width prediction was conducted in two distinct modes: "mass-inclusive", which treats mass as a predictive feature, and "mass-exclusive", which isolates the influence of quantum numbers alone.

The modeling strategy follows a three-level hierarchical approach designed to evaluate learning efficiency. First, baryon and meson subsets were modeled separately to capture family-specific regularities within homogeneous data. Second, these were combined into a shared feature space to establish a unified hadronic model. Finally, a direct comparison between these levels was conducted to determine whether predictive performance benefits more from subclass specialization or from a common feature space that captures dominant trends across both families.

To mitigate distribution shifts between the learning and evaluation phases, the datasets were partitioned using a balanced splitting strategy. The combined mass dataset was divided into train/validation/test (284/61/61), while the width dataset followed a 261/56/56 split, ensuring that the baryon-meson ratio remained consistent across all sets.

Exploratory analysis of the target variables revealed that while mass values are broadly distributed, widths exhibit a strongly right-skewed pattern with many states approaching zero. To address this numerical imbalance, width prediction was treated using both the original linear scale and a log(1+Γ) transformation. This transformation was crucial for the learning process, as it reduces the dominance of high-width outliers while preserving the subtle contributions of narrow states, resulting in a more balanced and stable target space.

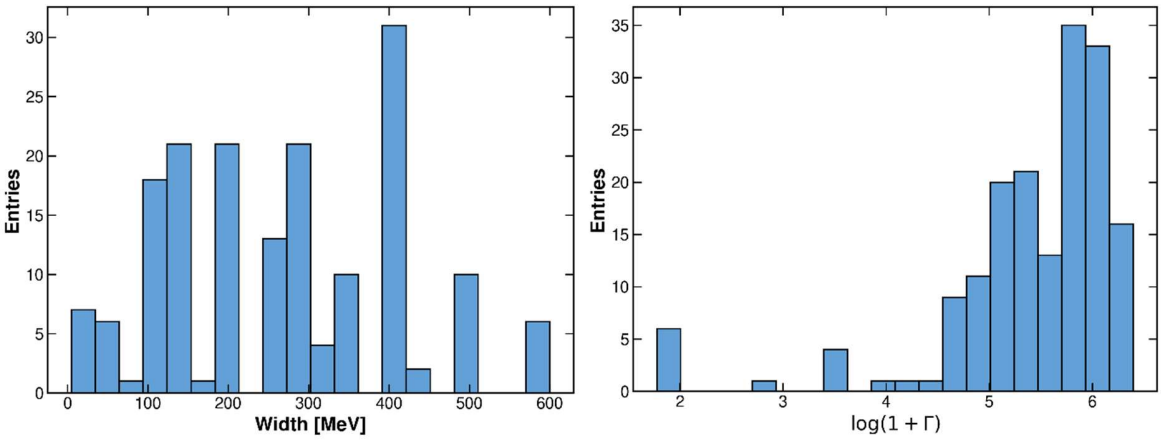

Figure 1. Width and log(1+Γ) distributions of baryons. The raw baryon width distribution is strongly right-skewed, whereas the transformed distribution is more balanced and closer to Gaussian, improving learnability.

2.2 XGBoost-based Regression Approach

In this study, XGBoost was preferred as the primary learning algorithm for mass and width prediction. XGBoost is a powerful ensemble learning method based on the regularized gradient boosted decision trees approach, which demonstrates high success on tabular data while also providing interpretability. Developed by Chen and Guestrin (Chen & Guestrin, 2016) , this method sequentially combines tree-based weak learners to gradually reduce residual errors, offering advantages such as regularization, sensitivity to sparse data, and computational efficiency.

The XGBoost regression model is defined through an additive structure:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in \mathcal{F}$$

Where represents the input vector, represents the predicted output, and represents the weak learners in the form of decision trees. The model aims to reduce the errors remaining from previous stages by adding a new tree at each iteration. The objective function used is the sum of a loss term measuring data fit and a regularization term controlling model complexity:

$$\mathcal{L}^{(t)} = \sum_i l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t)$$

and

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2$$

In these expressions, denotes the number of leaves, denotes the leaf weights, and and represent regularization parameters determining the complexity penalty. This framework enables the effective learning of non-linear interactions while limiting overfitting.

There are several key reasons for choosing XGBoost for this study. First, the dataset is clearly tabular, and there are non-linear interactions between quark contents and quantum numbers. Tree-based methods can learn such relationships without explicitly designing feature transformations. Second, since there are missing observations in some variables in the dataset, the flexible structure of XGBoost toward missing values provides a practical advantage. Third, the method is conducive to generating feature importance analysis, which allows for a physical interpretation of which quark contents or quantum numbers play a more dominant role in mass and width prediction. The purpose of the present study is not to provide a broad benchmark comparison among multiple machine-learning algorithms, but rather to investigate whether an XGBoost-based framework can successfully model hadron mass and decay width while also offering physically interpretable feature rankings. Earlier meson-focused studies, particularly the work by Akan, have already shown that several machine-learning approaches can be applied to hadronic prediction problems(Akan, 2024). Building on that foundation, the present study adopts XGBoost as the primary regression model because it is especially well suited to structured tabular data, can effectively capture non-linear interactions among quark content and quantum numbers, and enables an interpretable analysis through feature-importance scores.

To ensure reproducibility, the main hyperparameters of the XGBoost regressor were fixed explicitly before final testing. Hyperparameter values were selected through validation-guided empirical tuning based on preliminary training runs, while the test set was reserved strictly for final performance evaluation. This procedure was adopted to avoid information leakage from the test set into model development. Since parameters such as n_estimators, max_depth, learning_rate, subsample, and colsample_bytree strongly influence model flexibility and generalization, the final configurations are reported in Table 3. To preserve comparability across baryon, meson, and full-hadron analyses, the same fixed configuration was used within the corresponding mass and width modeling setups.

Table 3. XGBoost hyperparameter configuration. Fixed hyperparameter settings used in the mass and width prediction models. The table reports the fixed XGBoost hyperparameters adopted in the present study. These settings were determined through validation-guided preliminary tuning and then kept constant during final evaluation to ensure reproducibility and comparability across baryon, meson, and unified hadron analyses.

hyperparameter	mass_model	width_with_mass	width_without_mass
n_estimators	600	700	700
max_depth	4	4	4
learning_rate	0.03	0.05	0.05
subsample	0.9	0.9	0.9
colsample_bytree	0.9	0.9	0.9
objective	reg:squarederror	reg:squarederror	reg:squarederror
random_state	42	42	42
n_jobs	2	2	2

In this study, XGBoost was used for four main problem classes: (i) baryon mass prediction, (ii) baryon width prediction, (iii) meson mass prediction, and (iv) meson width prediction. Additionally, the following tasks were performed on the unified dataset: (v) hadron mass prediction, (vi) hadron width prediction using the mass feature, and (vii) hadron width prediction without using the mass feature. In this way, both subclass-based and holistic hadron modeling were compared under the same method family.

In mass prediction, the target variable was defined directly as mass in MeV. For width prediction, two different approaches were followed. The first is the model where the mass variable is included in the inputs, accepting that mass may be physically related to the decay width. The second is the model proceeding only from quark content and quantum numbers, i.e., excluding mass information. These two setups were designed to directly test to what extent width prediction depends on the hadron mass.

Additionally, two model variants, referred to as real only and augmented, were considered in this study. The real only variant denotes the baseline models trained exclusively on experimentally reported observations, whereas the augmented variant includes synthetic samples generated during preprocessing by applying Gaussian-noise perturbations to the original data. The purpose of this comparison is to assess whether controlled data augmentation can improve model generalization relative to the baseline configuration. In this section, however, the emphasis remains on the core XGBoost framework that defines the overall behavior of the model family; detailed comparisons between these variants are presented and interpreted in the Results and Discussion section.

2.3 Analysis Strategy and Evaluation Metrics

The analysis workflow of the study was structured in three main stages. The first stage is exploratory data analysis (EDA). Within this scope, the numerical distributions of hadron, baryon, and meson samples; mass and width histograms; target variable distributions by particle type; and specifically, behavior after logarithmic transformation for width were examined. This step is necessary to determine the balance of the dataset, the spread of the target variables, and potential outliers.

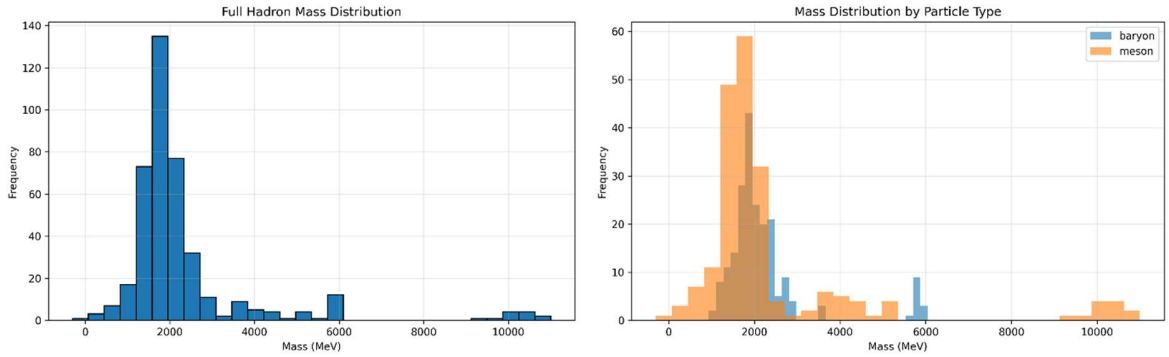

Figure 2. Mass Distribution of Full Hadrons (left) and by Particle Type EDA plots (right) visualizing the spread of mass values, showing that while most hadrons cluster in the 1–2.5 GeV band, mesons exhibit a wider spread in low and high-mass regions compared to baryons.

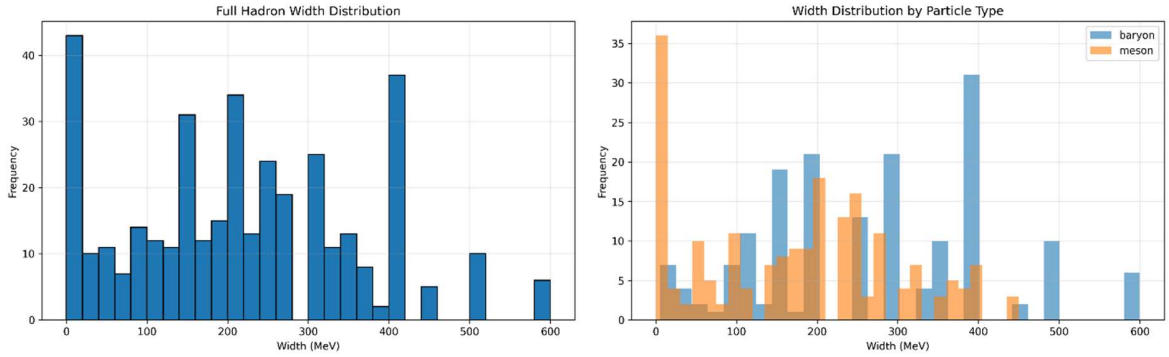

Figure 3. Width Distribution of Full Hadrons (left) and by Particle Type (right). Visualization of decay widths for the total dataset and subsets, indicating the skewed nature of width variables and the higher representative power of mesons in the width variable compared to baryons.

The second stage is relationship analysis. Here, linear trends between quark contents, quantum numbers, mass, and width were visualized using correlation matrices. The relationships between the target variable and features were evaluated in the mass task, and the correlation structure for both mass-inclusive and mass-exclusive scenarios was evaluated in the width task. This analysis was conducted to see which variables carry stronger links with the target magnitudes before modeling and to support the physical justification for feature selection.

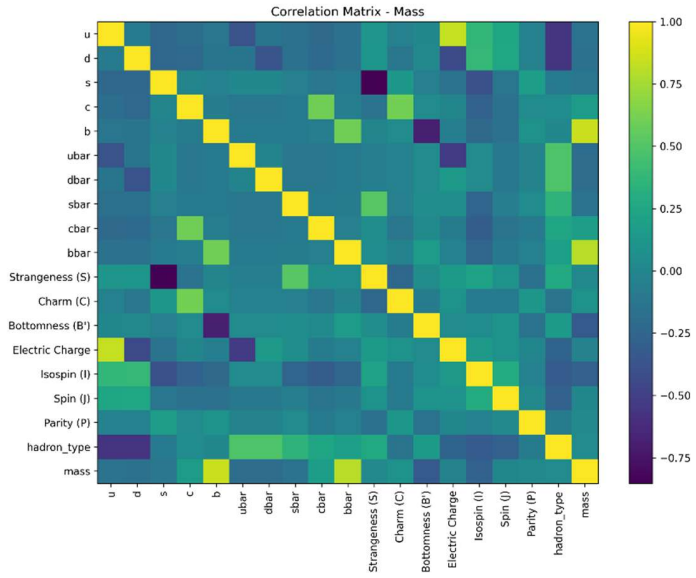

Figure 4.Correlation Matrix of Full Hadrons for Mass Variable. Matrix displaying the linear relationships between inputs and target mass, highlighting the strong positive correlation between mass and heavy flavor components like bottom and charm quarks.

The third stage is the quantitative and visual evaluation of prediction performance. In this context, Coefficient of Determination (R^2), Mean Absolute Error (MAE) , and Root Mean Square Error (RMSE) values were calculated on the validation and test sets for each model. These metrics collectively indicate the model's variance-explaining power; specifically, MAE provides a direct measure of the average error magnitude, while RMSE, due to its squared nature, offers higher sensitivity to large errors. Together, these three criteria ensure a balanced assessment of both the general fit and the magnitude of deviations. In addition to performance metrics, the following graphical diagnostic tools were used to examine model behavior in detail:

- True vs Predicted scatter plots: To show the proximity of predictions to the ideal line,
- Residual Histogram: To examine the centering and symmetry of the error distribution,
- Residual vs Predicted plots: To determine if there is a systematic bias depending on the prediction level,
- Feature Importance plots: To show which inputs XGBoost assigns more weight to,
- Top 10 Test Errors tables: To allow for the individual physical examination of particles with the largest errors.

The study was structured to answer not only "which model is better?" but also "which physical variables are decisive?", "in which regions do the error increase?", and "to what extent does mass information contribute to width prediction?". Thus, the method section allows the results to be addressed not just through numerical performance, but also in terms of physical interpretability and model behavior.

3. RESULTS AND DISCUSSION

3.1 Hadrons Mass Prediction

The XGBoost-based mass prediction model established on the combined hadron dataset demonstrated a high level of explanatory power by learning both baryon and meson samples within the same feature space. The base model, trained with real observations, achieved, R^2=0.874, MAE=263.52 MeV, and RMSE=372.63 MeV on the validation set; while on the test set, it obtained R^2=0.964, MAE=313.03 MeV, and RMSE=422.23 MeV. These results indicate that the model successfully captured the non-linear relationships between hadron mass and input variables. While the higher value in the test set compared to the validation set is noteworthy, this situation is likely due to differences in sample composition arising from the data split. Specifically, the limited size of the dataset and the small number of samples in the high-mass region can produce varying error profiles across the validation and test subsets.

Table 3. Full Hadron Mass Prediction Performance (Validation and Test Results). This table compares the predictive performance of the real_only and augmented model variants on validation and test sets. The results demonstrate that the real_only variant, trained exclusively on experimental observations, provides superior generalization compared to the augmented approach. Metrics such R^2, MAE, and RMSE confirm that data augmentation did not offer a performance advantage for this specific mass prediction task.

dataset	model_variant	R2	MAE	RMSE
validation	real_only	0.957202	251.385374	353.288950
test	real_only	0.941493	259.481589	356.773293
validation	augmented	0.930760	281.881722	441.388823
test	augmented	0.946217	269.843771	340.337458

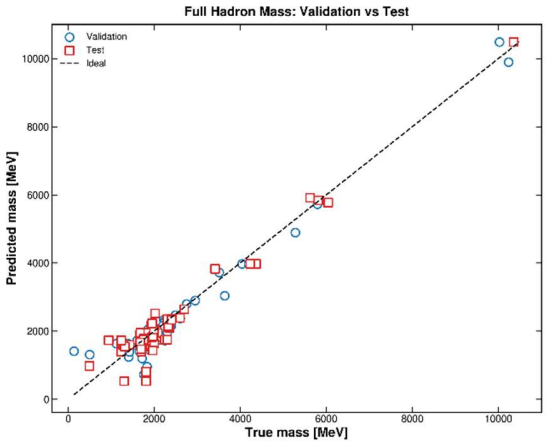

Figure 5. Hadron Mass analysis Validation and Test comparison. Scatter plot showing the proximity of predicted mass values to the ideal line, demonstrating the model's success in capturing non-linear mass patterns across the full hadron spectrum.

When comparing model variants, the augmented approach involving data augmentation did not provide a performance advantage over the real only base model. For the augmented variant, R^2=0.856, MAE=275.78 MeV, and RMSE=397.98 MeV were obtained on the validation set. This suggests that the base model, which preserves the distribution of original physical observations, offers a more balanced generalization for this specific problem. Therefore, interpretations regarding hadron mass are primarily based on the real only model.

Table 4. Full Hadron Mass Prediction Performance (Model Variants). Comparison of R2, MAE, and RMSE metrics between the real_only baseline and the augmented model, showing that preserving original physical observations provides better generalization for mass prediction.

model_variant	R2	MAE	RMSE
real_only	0.957202	251.385374	353.288950
augmented	0.930760	281.881722	441.388823

Sub-dataset-based test results reveal the behavior of the unified model on baryon and meson samples in more detail. In the baryon test subset, R^2=0.923, MAE=227.26 MeV, and RMSE=302.95 MeV were found; while in the meson test subset, R^2=0.964, MAE=401.66 MeV, and RMSE=517.37 MeV were obtained. This situation presents a seemingly contradictory but physically meaningful result. The larger variance within the meson subset naturally leads to a higher proportion of explained variance or R^2. At the same time the wider mass scale inherent in this group results in an increase in absolute error magnitudes. In other words, while the model captures a relatively better mass ranking for mesons, the error magnitude in MeV is higher than for baryons due to the broader and multi-modal nature of the meson mass spectrum. In the baryon subset, error magnitudes remained more limited because of the narrower mass range.

Table 5. Sub-dataset-based Test Results (Baryon and Meson). Detailed evaluation of the unified model on specific subsets, illustrating that while the model captures mass rankings well for mesons, absolute error magnitudes are higher due to the broader meson mass scale.

subset	R2	MAE	RMSE
baryon_test_subset	0.950610	195.977726	269.103445
meson_test_subset	0.936664	325.102247	428.936730

Exploratory graphics support this interpretation. The mass distribution of all hadrons is significantly right-skewed, with most observations clustered in the 1–2.5 GeV band; in contrast, a smaller number of heavy hadrons form a long tail extending into several GeV regions up to approximately 10 GeV. When examined by particle type, mesons exhibit a wider spread in both the low-mass and very high-mass regions, whereas baryons are mostly concentrated in the medium-mass region. This heterogeneous structure requires the unified model to learn different physical scales simultaneously and directly influences the interpretation of feature importance.

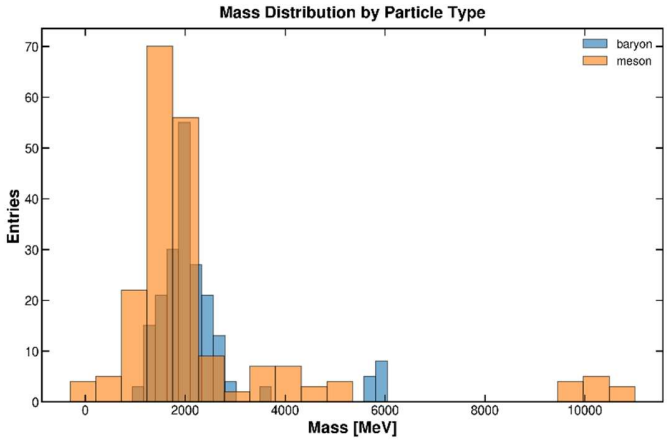

Figure 6. Mass Distribution by Particle Type. This histogram compares the mass distributions of baryons and mesons within the dataset. It shows that mesons exhibit a wider spread in both low and very high-mass regions, while baryons are primarily concentrated in the medium-mass range (1–2.5 GeV).

XGBoost feature importance results shown in Figure 7 clearly show that heavy quark content plays a dominant role in determining hadron mass. Specifically, the number of b quarks stands out as the variable with the highest importance, followed by, Bottomness (B'), c, and hadron type . In contrast, the contributions of light quark variables, electric charge, isospin, spin, and parity remain much more limited. For the unified hadron model, the limited but non-zero contribution of the hadron type variable indicates that baryon and meson classes carry some systematic differences in mass formation, but the primary decisive factor remains the heavy flavor content. Similarly, the correlation matrix reveals that mass has a strong positive relationship particularly with and content, while charm and other flavor components provide secondary but significant contributions. This result is consistent with the direct effect of heavy quarks on hadron mass.

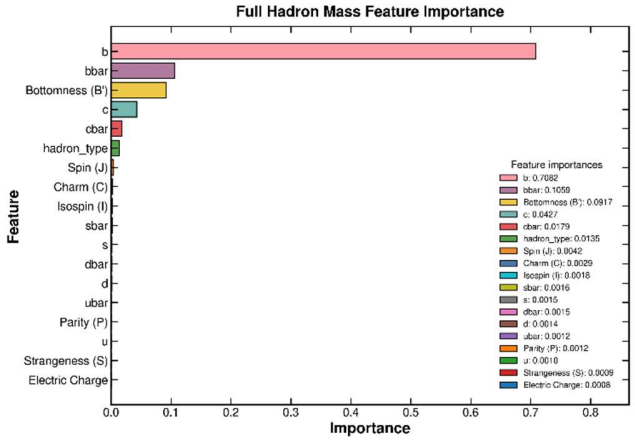

Figure 7. Full Hadron Mass Feature Importance. XGBoost ranking of input variables, identifying heavy quark content (specifically b quarks) as the dominant factor in determining hadron mass.

Scatter plots comparing true and predicted masses given in Figure 8 show that the model produces predictions close to the ideal line for the vast majority of samples. The general trend has been successfully captured, especially for high-mass hadrons. Conversely, residual distributions indicate that model errors are not completely randomly distributed and that significant deviations occur in some specific resonance states. While the residual histogram is concentrated around zero, there are long tails in both directions. Residual plots against predictions reveal that the limited number of observations in the very high-mass region and certain resonances around 1–2 GeV pose more difficulty for the model. This suggests that quark content and basic quantum numbers alone may not always be sufficient to fully represent the fine structures across the entire hadron mass spectrum.

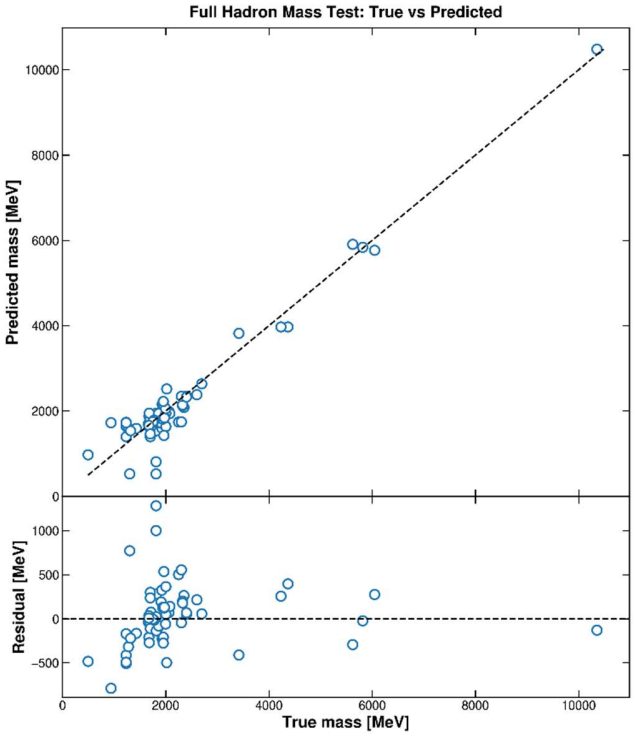

Figure 8. Hadron Mass True vs. Predicted (top panel) and Residual (bottom panel) Plot. This figure displays a scatter plot comparing true versus predicted hadron masses alongside their corresponding residuals. The upper panel shows that predictions are closely aligned with the ideal diagonal line, while the lower panel illustrates the distribution of errors (residuals) across the mass spectrum.

An examination of the top 10 samples with the largest test errors, given in the Table 7, strengthens this interpretation. The highest absolute errors largely occurred in meson states. Specifically, meson states such as $\Upsilon(1S)$, $\varphi(1020)$, $\pi(1300)^-$, $\pi(1800)^+$, $\psi(4660)$, $f_4(2050)$ and $a_0(1450)^-$ are at the top of the list. On the baryon side, the n (neutron), $N(2570)5/2^-$ and $\Delta(1232)3/2^+$ states are notable. First, the growth of absolute error in very high-mass states does not always mean relative failure; for example, while the error for is large in absolute terms, the total mass scale is also quite high. Second, the large deviations seen in excited meson and baryon resonances in the medium-mass region suggest that fine spectral distinctions between states with the same quark content cannot be fully resolved with the current input variables alone.

Table 6. Top 10 Samples with Largest Mass Test Errors. Identification of specific particles with the highest prediction deviations, such as excited meson and baryon resonances, suggesting that current inputs may not fully resolve fine spectral distinctions.

Name	hadron_type	true_mass	predicted_mass	absolute_error	residual
π(1800)−	meson	1812.000000	529.015259	1282.984741	1282.984741
π(1800)+	meson	1812.000000	810.105896	1001.894104	1001.894104
n (neutron)	baryon	939.565000	1727.512329	787.947329	-787.947329
π(1300)−	meson	1300.000000	529.015259	770.984741	770.984741
N(2300) 1/2	baryon	2300.000000	1745.381592	554.618408	554.618408
h1(1965)	meson	1965.000000	1429.554077	535.445923	535.445923
Δ(1232) 3/2	baryon	1232.000000	1738.596313	506.596313	-506.596313
K2(2250)0	meson	2247.000000	1743.938477	503.061523	503.061523
f4(2050)	meson	2018.000000	2516.068359	498.068359	-498.068359
Δ(1232) 3/2	baryon	1232.000000	1724.526001	492.526001	-492.526001

3.2 Hadrons Width Prediction

Hadron decay width is a more complex physical quantity compared to mass, closely related not only to internal quark content and conserved quantum numbers but also to accessible decay channels, threshold effects, and resonance structure. Therefore, width prediction constitutes a more challenging regression problem. This study examined width prediction under two different modeling setups: first, where width is predicted solely from quark content and quantum numbers, and second, where hadron mass is included as an additional input variable. This allowed for a direct test of the role of mass in width prediction.

Table 7. Comparison of Width Modeling Setups. Quantitative summary comparing width prediction without mass information vs. including mass, highlighting the significant performance boost when mass is used as a predictor.

dataset	experiment	R2	MAE	RMSE
validation	width_without_mass	0.390003	73.681350	106.228736
test	width_without_mass	0.672490	65.624848	96.229149
validation	width_with_mass	0.493721	55.880688	96.777272
test	width_with_mass	0.879525	40.213011	58.363718

The model established on the combined hadron dataset without mass information produced R^2=0.309, MAE=87.03 MeV, and RMSE=113.83 MeV on the validation set; on the test set, performance further declined to R^2=0.146, MAE=87.33 MeV, and RMSE=122.39 MeV. This indicates that hadron width cannot be modeled with sufficient accuracy using only quark content and basic quantum numbers with this model (XGBoost). The limited value in the test set suggests the model explains only a small fraction of the observed variance. This is physically expected, as decay width is directly connected to accessible decay channels and their dynamic dominance.

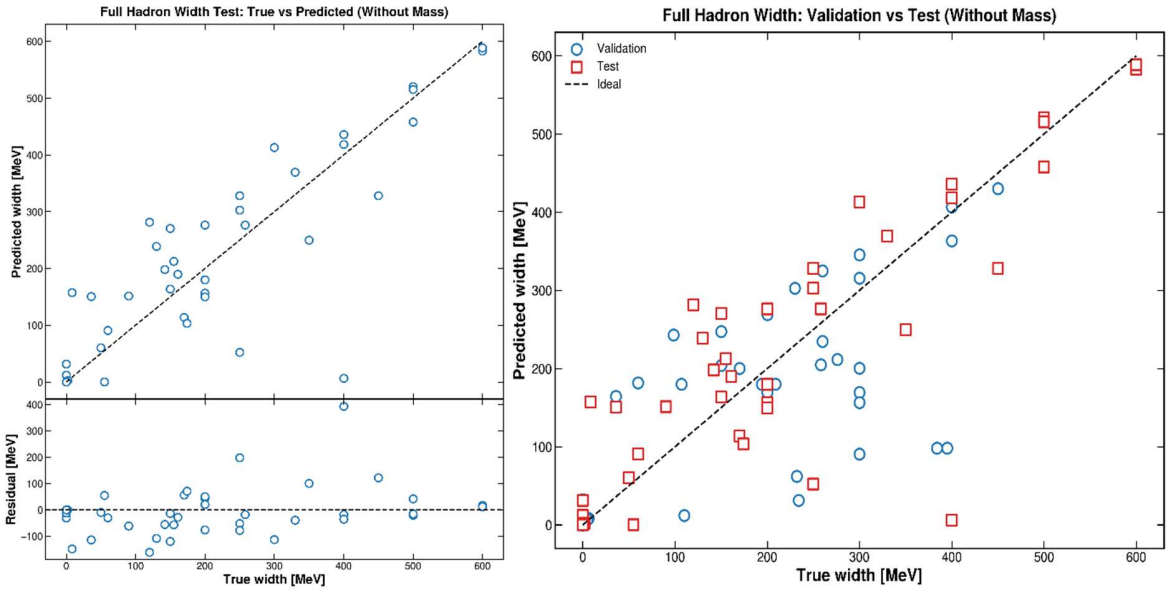

Figure 9. The limited predictive power of the model without mass input is visualized in figures, where both the parity relation and the residual structure reveal substantial deviations.

Conversely, the width model using hadron mass as an additional input variable exhibited significant improvement. This setup achieved R^2=0.804, MAE=36.89 MeV, and RMSE=60.55 MeV on the validation set, and R^2=0.812, MAE=36.95 MeV, and RMSE=57.38 MeV on the test set. This performance boost clearly demonstrates a strong and systematic relationship between hadron width and mass. The primary determinant for width prediction success is the mass information, which defines the kinematic decay regime. The strong gain obtained in the mass-included setup is physically expected. Within the resonance picture, the decay width depends not only on internal quantum numbers but also on the resonance mass. This mass variable is crucial because it controls the available phase space and the opening of decay channels. Consequently, the result is consistent with Breit–Wigner-type resonance behavior. The substantial performance gain obtained when mass is added is highly significant. It shows that the model does not merely exploit a numerical correlation. Instead, it successfully learns the physically expected connection between resonance mass, decay kinematics, and total width.

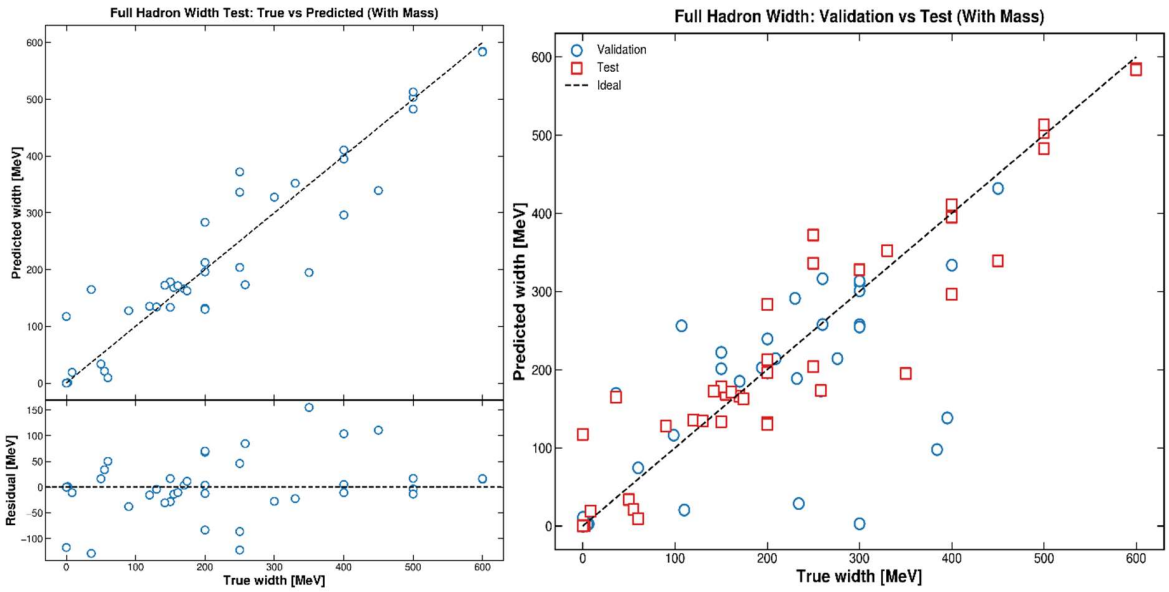

Figure 10. The substantial improvement obtained after including mass as an input variable is clearly seen in figure, where the predictions become much more tightly aligned with the ideal trend.

The combined hadron width dataset consists of 373 samples (172 baryons, 201 mesons). The train/validation/test split followed a 261/56/56 ratio, maintaining the baryon-meson distribution in each subset.

Table 8. Width Dataset Split Statistics. This table provides a statistical summary of the dataset used for hadron width prediction, divided into training, validation, and test sets. It details the sample counts for baryons and mesons within each split, alongside key distribution metrics such as minimum, median, maximum, and quartile width values (in MeV).

split	count	baryon_count	meson_count	width_min	width_q1	width_median	width_q3	width_max
train	182	84	98	0.000000	103.750000	200.000000	318.250000	600.000000
validation	39	18	21	0.000000	48.000000	200.000000	300.000000	450.000000
test	40	18	22	0.000000	112.500000	200.000000	335.000000	600.000000

XGBoost feature importance plots show that in the model without mass, importance values are scattered and relatively weak across hadron type, flavor content, spin, and parity. This suggests the model relies on multiple weak signals. When mass is included, it becomes the overwhelmingly dominant variable, confirming that mass is not just an auxiliary feature but the primary explanatory variable for hadron width.

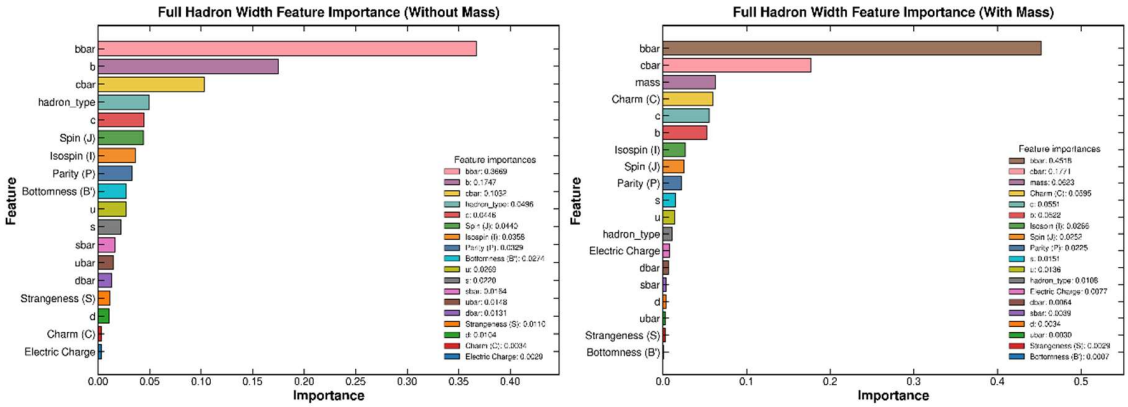

Figure 11. This contrast is directly reflected in the feature-importance distributions shown in figures: without mass, the model relies on several weak signals, whereas with mass, the mass variable becomes clearly dominant.

3.3 Comparison of Hadron-level Results with Baryon and Meson Subsets

In addition to evaluating the unified hadron model, the present study also compares its behavior with the corresponding baryon-only and meson-only models. This comparison is important because it reveals whether the predictive structure is governed mainly by global hadronic regularities or by subclass-specific patterns. Accordingly, the following subsection contrasts baryon, meson, and full-hadron results for mass prediction, width prediction without mass information, and width prediction with mass information, thereby providing a more complete physical and methodological interpretation of model performance.

3.3.1. Comparison in Terms of Mass Prediction

When the mass prediction results are examined together, it is evident that high success was achieved in both baryon and meson subsets. The baryon model produced R^2=0.942 and R^2=0.904 on the validation and test sets respectively, while the meson model reached R^2=0.946 and R^2=0.959 on the same sets. The unified hadron model yielded R^2=0.874 on the validation set and R^2=0.964 on the test set. These results confirm that the mass problem can be strongly learned by XGBoost and that quark content and quantum numbers possess high explanatory power for hadron mass.

When examined at the subset level, the meson model produced the highest R^2 value particularly on the test set. This suggests that the updated meson data representation offers a more discriminative feature space for the model. The explicit encoding of the quark-antiquark structure in mesons creates distinct separations in the mass scale, especially for heavy mesons containing charm and bottom quarks, facilitating the model's ability to learn these distinctions. Although the baryon model's test performance is also very high, it is slightly lower compared to mesons. A likely reason for this is that different resonances and excited states sharing the same quark content in baryons cluster in closer mass regions, making discrimination more difficult.

The fact that the unified hadron model produces performance close to the meson model on the test set is noteworthy. However, the lower R^2 on the validation set compared to the subset models indicates that a single model faces a more heterogeneous problem when learning both baryon and meson spectra simultaneously. In other words, while unified hadron-level modeling can successfully capture the overall mass pattern, it cannot represent the finer structures within each subclass as efficiently as separate models.

The MAE and RMSE values further support this picture. The baryon mass model yielded MAE=239.16 MeV and RMSE=324.14 MeV on the test set, while the meson model produced MAE=406.30 MeV and RMSE=532.97 MeV. The hadron model lies between these two extremes with MAE=313.03 MeV and RMSE=422.23 MeV. This indicates that the unified model strikes a balance between hadron classes, but cannot fully match the specialized performance of subset models in terms of absolute error levels, particularly due to the wide mass range of mesons and the denser medium-mass structure of baryons.

Table 9. Mass Prediction Performance Summary. This table compares R2, MAE, and RMSE metrics for mass prediction across baryon, meson, and unified hadron datasets. It highlights that high accuracy is achieved in all categories, with the unified model reaching a test R2 of 0.964.

dataset	validation_R2	test_R2	validation_MAE	test_MAE	validation_RMSE	test_RMSE
baryon	0.942000	0.904000	259.990308	239.158208	328.006164	324.141843
meson	0.946000	0.959000	295.257844	406.295979	393.815205	532.966588
hadron	0.874000	0.964000	263.521411	313.030590	372.630373	422.234162

3.3.2. Comparison in Terms of Width Prediction: Without Mass Information

Differences between subsets and the unified hadron model become considerably more pronounced in width prediction. In the model setup without mass information, the baryon model produced R^2=0.475 and R^2=0.469 on the validation and test sets respectively. In contrast, the meson model yielded only R^2=0.139 and R^2=0.203 for the same problem. The unified hadron model remained at R^2=0.309 on the validation set and R^2=0.146 on the test set.

These results suggest that explaining the width variable solely through quark content and quantum numbers is particularly difficult for mesons among the two hadron classes. The better performance of the baryon model compared to the meson model indicates that baryon widths carry a more regular and learnable pattern within the existing feature space. Since mesons with the same quark composition can have a large number of resonances, mixing configurations, and channel-dependent decay structures, width becomes increasingly distant from being a quantity directly derivable from structural variables alone.

The unified hadron model was found to occupy an intermediate position between baryons and mesons in this scenario. This is understandable, as the model simultaneously attempts to learn the relatively more regular width behavior seen in baryons on one hand, and the more complex and heterogeneous width structure of mesons on the other. The fact that the test set R^2 drops as low as 0.146 demonstrates that the unified model cannot adequately discriminate the width problem at the hadron level without mass information. In other words, when mass is excluded, the single unified model approach for width prediction at the hadron level is not as effective as subset-specific models.

Table 10. Width Prediction Performance Without Mass Information. This table compares R2, MAE, and RMSE metrics for width prediction across baryon, meson, and unified hadron datasets without using mass as an input. It highlights the limited explanatory power of structural variables alone, particularly for mesons and the unified model.

dataset	validation_R2	test_R2	validation_MAE	test_MAE	validation_RMSE	test_RMSE
baryon	0.475000	0.469000	78.963604	79.960552	107.913160	105.869986
meson	0.139000	0.203000	77.075511	76.043664	105.652088	110.331234
hadron	0.309000	0.146000	87.030508	87.330083	113.830090	122.394586

3.3.3. Comparison in Terms of Width Prediction: With Mass Information

Including the mass variable fundamentally changes the picture. The baryon model produced R^2=0.898 and R^2=0.922 on the validation and test sets respectively. In the meson model, the validation performance reached a very high level of R^2=0.941, while the test performance was R^2=0.748. The unified hadron model yielded R^2=0.804 and R^2=0.812 on the validation and test sets respectively.

The most important physical implication of these results is that mass is a central variable in the prediction of hadron width. Since width depends not only on the internal quark structure of the hadron but also on which decay channels are kinematically open, mass information adds a powerful explanatory dimension to the model. The dramatic improvement in the baryon model demonstrates this very clearly: the R^2 value rose from approximately 0.47 to 0.92. The same trend holds for the meson model; however, the gap between validation and test performance remains larger for mesons. This suggests that the meson width problem carries a more complex, more heterogeneous, and harder to generalize structure compared to baryons.

The unified hadron model, when mass is included, meaningfully captures the common pattern between baryons and mesons and reaches R^2=0.812 on the test set. This performance falls behind the baryon model but is close to, and slightly above, the meson model's test performance. The conclusion to be drawn here is that mass information establishes a strong enough physical link to make unified hadron-level modeling feasible; however, subset-specific specialized models, particularly for baryons, can still produce higher accuracy.

Table 11. Width Prediction Performance with Mass Information. This table compares R2, MAE, and RMSE metrics for width prediction across baryon, meson, and unified hadron datasets with mass included as an input. It highlights the dramatic improvement in predictive power across all categories when mass information is utilized.

dataset	validation_R2	test_R2	validation_MAE	test_MAE	validation_RMSE	test_RMSE
baryon	0.898000	0.922000	22.767252	19.825544	47.544889	40.692741
meson	0.941000	0.748000	19.688495	26.872123	27.715616	62.004393
hadron	0.804000	0.812000	36.885290	36.954480	60.545543	57.378470

3.3.4. Physical and Methodological Implications

- When the baryon, meson, and unified hadron results are evaluated together, three key conclusions can be drawn.
- Mass prediction is highly successful across all datasets, confirming the strong link between quark content and mass.
- Width prediction without mass is extremely challenging, especially for mesons, due to complex decay dynamics.

Mass is a dominant physical indicator for width, determining the decay phase space. Methodologically, unified models are useful for general trends, but subset-specific models offer higher precision and clearer physical interpretation.

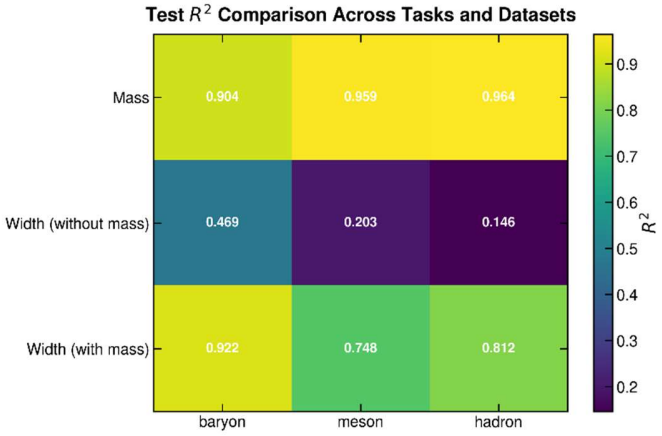

Figure 12. Test R^2 Comparison Across Tasks and Datasets. This heatmap summarizes the model performance R2 for mass and width prediction across different datasets. It visually demonstrates that while mass prediction is highly successful for all groups, width prediction is significantly more challenging and requires mass as an input variable to achieve high accuracy

4. CONCLUSION

In this study, the predictability of hadron mass and decay width was systematically investigated using an XGBoost-based machine learning approach on a hadron dataset composed of baryons and mesons. The model inputs were constructed from the quark content of hadrons together with quantum numbers such as isospin, spin, parity, electric charge, strangeness, charmness, and bottomness. The

analyses were carried out separately for the baryon, meson, and combined hadron datasets. For the width problem, two different scenarios were also considered: one in which hadron mass was included in the model and another in which it was excluded. In this way, both the explanatory power of structural variables and the role of mass in width prediction were examined in a data-driven manner.

The findings show that hadron mass can be predicted with high accuracy from quark content and quantum numbers. High R^2 values, reaching 0.904 for baryons, 0.959 for mesons, and 0.964 for the combined hadron dataset on the test set, were obtained for mass prediction, indicating that hadron mass has a strong and learnable relationship with the input variables. In particular, heavy-flavor content was found to play a dominant role in mass prediction, as clearly reflected in the feature-importance analysis and correlation structures. This result is consistent with the physical expectation that hadron mass is largely determined by quark composition.

In contrast, hadron width prediction was considerably more challenging than mass prediction. Using only quark content and quantum numbers led to limited performance, especially for mesons, showing that width is not governed solely by static structural variables. Including hadron mass produced a clear and systematic improvement. This is physically expected, since in the resonance picture the decay width depends on the resonance mass through the available phase space and the kinematic opening of decay channels, in a manner consistent with Breit–Wigner-type behavior. The gain in performance therefore indicates that the model captures this expected physical dependence rather than merely exploiting numerical correlations. Data augmentation, on the other hand, did not improve the results, suggesting that preserving the original physical distribution is more effective for this task. Comparisons across subsets also provided important insights. The models constructed for baryons were found to be more stable and to exhibit better generalization performance than the meson models, particularly for width prediction. For mesons, mass prediction was very strong, whereas the width problem displayed a more heterogeneous and more difficult-to-learn structure. Although the combined hadron model was able to capture the overall trends successfully, it became clear that, especially in width prediction, baryon- and meson-specific details are more difficult to represent within a single unified model. Therefore, while the combined modeling approach is useful for examining global trends, subclass-based modeling appears to be more suitable for more precise predictions.

Another important outcome of this study is that XGBoost provides not only high predictive performance but also an interpretable modeling framework. The feature-importance plots revealed which quark contents and quantum numbers were more dominant in determining the target variables, while the residual analyses and true-versus-predicted comparisons showed in which physical regions the model performed more successfully and where it remained limited. In this respect, the study treats machine learning not merely as a numerical prediction tool, but as an analytical framework capable of making visible the patterns that govern hadron properties.

Table 12. Consolidated Summary of Predictive Performance. This table provides a comprehensive overview of the predictive performance metrics (R2, MAE, and RMSE) for all experimental tasks across the baryon, meson, and unified hadron datasets.

dataset	task	validation_R2	test_R2	validation_MAE	test_MAE	validation_RMSE	test_RMSE
baryon	mass	0.942083	0.904073	259.990308	239.158208	328.006164	324.141843
meson	mass	0.946462	0.959148	295.257844	406.295979	393.815205	532.966588
hadron	mass	0.873874	0.963889	263.521411	313.030590	372.630373	422.234162
baryon	width_without_mass	0.475343	0.469002	78.963604	79.960552	107.913160	105.869986
meson	width_without_mass	0.139182	0.202713	77.075511	76.043664	105.652088	110.331234
hadron	width_without_mass	0.308543	0.146089	87.030508	87.330083	113.830090	122.394586
baryon	width_with_mass	0.898156	0.921552	22.767252	19.825544	47.544889	40.692741
meson	width_with_mass	0.940761	0.748196	19.688495	26.872123	27.715616	62.004393
hadron	width_with_mass	0.804379	0.812334	36.885290	36.954480	60.545543	57.378470

In conclusion, this study demonstrates that the XGBoost-based data-driven approach is a powerful method for hadron spectroscopy, particularly for mass prediction, and that it becomes effective for width prediction when supported by appropriate physical variables. The results suggest that machine learning can provide a complementary contribution to both theoretical and experimental studies in hadron physics, especially by offering useful predictions for hadron states where data are limited, uncertainties are high, or physical properties have not yet been fully established. In future work, several directions appear particularly promising. Given the observed difficulty in generalizing meson width prediction, where validation and test R^2 values diverged notably, the development of resonance-aware feature representations that explicitly encode decay channel information could substantially improve performance. Additionally, probabilistic models incorporating uncertainty estimation would be valuable for hadron states where experimental measurements are sparse or conflicting. Expanding the dataset to include exotic hadron candidates such as tetraquarks and pentaquarks, and benchmarking

XGBoost against deep learning approaches on the same unified feature space, would further clarify the boundaries and advantages of the data-driven approach in hadron spectroscopy.

ACKNOWLEDGMENT

Special thanks to High Energy and Quantum Technologies (YEKUT) Laboratory, within the Department of Physics at the Faculty of Science and Letters, Yozgat Bozok University, for providing the computational logistics and infrastructure for this study

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

Akan, T. (2024). Exploring machine learning models for predicting meson mass and width. Physica Scripta, 99(2), 025203. https://doi.org/10.1088/1402-4896/ad1902

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785

Malekhosseini, M., Rostami, S., Olamaei, A. R., Ostovar, R., & Azizi, K. (2024). Meson mass and width: Deep learning approach. Physical Review D, 110, 054011. https://doi.org/10.1103/PhysRevD.110.054011

Navas, S., Amsler, C., Gutsche, T., Hanhart, C., Hernández-Rey, J. J., Lourenço, C., Masoni, A., Mikhasenko, M., Mitchell, R. E., Patrignani, C., et al. (Particle Data Group). (2024). Review of particle physics. Physical Review D, 110, 030001. https://pdg.lbl.gov/

Nielsen, D. (2016). Tree boosting with XGBoost: Why does XGBoost win "every" machine learning competition? [Master's thesis, Norwegian University of Science and Technology]. NTNU Open. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2433761

Particle Data Group. (2020). 2020 review of particle physics: Interactive listings. https://pdg.web.cern.ch/pdg/2020/listings/contents_listings.html

Tong, X., Feng, W., Xu, W., & Li, X. Q. (2025). Meson properties and symmetry emergence based on the deep neural network. Physical Review D, 111, 014023. https://doi.org/10.1103/PhysRevD.111.014023

Workman, R. L., Burkert, V. D., Crede, V., Klempt, E., Thoma, U., Tiator, L., Agashe, K., Aielli, G., Allanach, B. C., Amsler, C., et al. (Particle Data Group). (2022). Review of particle physics. Progress of Theoretical and Experimental Physics, 2022, 083C01. https://doi.org/10.1093/ptep/ptac097

Yasser, A. M., Nahool, T. A., Anwar, M., & Mansour, N. A. (2020). A new machine learning approach for predicting the spectra of meson bound states. International Journal of Modern Physics A, 35(32), 2050214. https://doi.org/10.1142/S0217751X2050214X