HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 4 (2024)

Leveraging Machine Learning for Accurate Star Formation Rate Predictions with MAGPHYS

Vishnu Parthasarathy[1] and Victoria Lloyd[#]

[1]Valley Christian High School, USA
[#]Advisor

ABSTRACT

Star formation rates (SFRs) are pivotal for understanding the growth of stars, galaxies, and the universe. Understanding SFR is essential for insights into galaxy evolution, stellar populations, cosmology, and interstellar dynamics. SFR analysis is well-suited to machine learning due to its complexity and volume of data. In our study, we utilized machine learning models on a dataset containing various factors such as gas luminosity, star formation timescale, and metallicity to predict SFRs. Our models included Linear Regression, Lasso Regression, and a neural network. Both Linear Regression and Lasso Regression yielded low mean squared error values, with the neural network achieving even lower values, demonstrating the superior performance of deep learning in determining SFRs. Additionally, we assessed feature importance for the Linear and Lasso Regression models, identifying which factors most significantly influence SFR predictions. From our analysis, we concluded that the aforementioned factors are crucial for accurately identifying SFRs in a galaxy, as our results showed that machine learning can predict SFRs with a mean squared error of 0.000939 and R-squared of 0.4808 based on galactic properties. Furthermore, we used graphs to illustrate the relationships between SFRs and different galactic properties, providing visual evidence of these connections. Our findings underscore the potential of machine learning in astrophysical research, particularly in predicting and understanding the intricate processes that govern star formation in various galactic environments. This approach can significantly enhance our comprehension of the universe's evolution.

Introduction

The formation of stars is a fundamental process that shapes the evolution of galaxies and the universe as a whole. Understanding the star formation rate (SFR) is crucial for astrophysics, as it provides insights into the lifecycle of galaxies, the interstellar medium (ISM), and the distribution of matter in the cosmos. Traditionally, determining the SFR involves complex techniques and theoretical models, which can be time-consuming and require extensive human expertise. However, recent advancements in artificial intelligence (AI) and machine learning (ML) offer a novel approach to tackling this challenge.

From Calzetti et al. 2007, we learn that there are many factors affecting star formation rate, including dust mass, dust attenuation, luminosity, and different factors of the ISM. From Moustakas et al. 2005, we learn of other factors influencing star formation rates, such as metallicity, ionization, and redshift. Finally, McKee et al. 2007 shows how turbulence, magnetic fields, and dust and gas density also affect star formation rates. This study will investigate some of these factors using machine learning.

AI, particularly deep learning, has shown remarkable success in various fields by leveraging large datasets to uncover patterns and make predictions. In astrophysics, the application of AI is still in its initial stages but is rapidly gaining traction. The potential of AI to process vast amounts of data, identify correlations, and predict outcomes with high accuracy makes it a valuable tool for determining the SFRs. By analyzing data derived from the spectral energy distributions (SEDs) of galaxies, AI can speed up the normally tedious process of estimating SFR.

This research paper explores the development and implementation of an AI-based framework to estimate the SFR from astronomical data. We employ a Linear Regression model, a Lasso Regression model, and a Neural Network, trained on a dataset containing galaxy information and corresponding SFRs derived from established methods, to predict the SFR directly from galactic information. Our approach aims to bridge the gap between traditional methods and modern computational techniques, providing a scalable and efficient solution for astronomers.

V. Bonjean et al. 2019 trained a Random Forest model on SDSS-DR8, using redshifts, WISE Luminosities, WISE colors in the near-IR, and spectra-extracted SFR and Stellar Mass. This model was trained on redshift (z) levels of $0.01 < z < 0.3$, and obtained a standard deviation of 0.38 dex for SFR and 0.16 dex for Stellar Mass. This method accurately predicted star formation rates without the need for complex modeling and only required the previously mentioned inputs. However, for galaxies with a redshift outside of this range, the model proved inaccurate.

Surana et al. 2020 trained a deep learning model on the GAMA Panchromatic Data Release to emulate the MagPhys model. Multiband flux measurements and the redshift were used to predict galactic properties such as stellar mass, dust luminosity, and star formation rates. This model produced standard deviation of 0.0577 for stellar mass prediction, 0.1643 for star formation rate prediction, and 0.1143 for dust luminosity prediction. This model proved much faster than the current MagPhys model, taking 0.03% of the time to predict factors for the same number of galaxies.

Background

Our models will look to use the derived MagPhys results as a labeled dataset to predict star formation rates. Our research will not use spectroscopic measurements, flux measurements, or redshift, and will instead use the derived characteristics of galaxies to predict the star formation rate. This will indicate how accurate machine learning with the MagPhys model is at discovering the importance of factors affecting star formation rates. It will also hold for a larger region of redshift than $0.01 < z < 0.3$, as the MagPhys DMU from GAMA is determined for galaxies in $z > 0.001$. By expanding the redshift range, our study aims to better understand star formation over more of the universe's history. We will examine various galactic features such as metallicity, gas, and dust mass, and the galaxy's environment to see how they influence star formation rates.

Dataset

For this project, we used the Galaxy and Mass Assembly (GAMA) Survey (Driver et al. 2022). This dataset is a spectroscopic survey of ~300000 galaxies, building on previous spectroscopic surveys such as the Millenium Galaxy Catalogue (J. Liske et al. 2003), the 2dFGRS (Colless M.M. 1999), and the Sloan Digital Sky Survey DR8 - (Eisenstein et al. 2011). The survey has a number of Data Management Units (DMUs), from which MagPhysv06 from GAMA II was selected. MagPhys (da Cunha et al. n.d.) is a model used to analyze spectral energy distributions (SEDs) to determine a number of characteristics, including but not limited to the age of the oldest stars in the galaxy, the star formation timescale, and the metallicity of the galaxy. GAMA's MagPhys DMU contains only the galaxies with redshift greater than 0.001, and thus contains data for 197494 total galaxies.

Due to computational limitations and network issues, we were only able to download 30000 galaxies from the MagPhys DMU, and therefore were only able to train our models on 30000 datapoints. Each galaxy had 18 parameters, shown in Table 1.

Table 1. Each galactic property derived from the MagPhys model and their definition.

Label:	Definition:

fmu(Opt)	fraction of total dust luminosity contributed by dust in the ambient (diffuse) ISM, optical
fmu(IR)	fraction of total dust luminosity contributed by dust in the ambient (diffuse) ISM, infrared
tform/yr	age of the oldest stars in the galaxy
gamma	star formation timescale (in Gyr)
Z/Z_o	metallicity Z
tauV	total V -band optical depth of the dust seen by young stars in their birth clouds
mu	fraction of tauV contributed by dust in the ambient (diffuse) ISM
$M*/M_{sun}$	Stellar mass
SFR(1e8)	Star formation rate, multiplied by 100 million
L_d/L_{sun}	Total stellar luminosity
ξ_c^{ISM}	fractional contribution by cold dust to the dust luminosity of the ambient ISM
$T_C^{BC/K}$	equilibrium temperature of warm dust in stellar birth clouds, in Kelvin
$T_C^{ISM/K}$	equilibrium temperature of cold dust in the ambient ISM, in Kelvin
ξ_{PAH}^{BC}	fractional contribution by Polycyclic Aromatic Hydrocarbon (PAHs) to the dust luminosity of stellar birth clouds
ξ_{MIR}^{BC}	fractional contribution by the hot mid-infrared continuum to the dust luminosity of stellar birth clouds
ξ_W^{BC}	fractional contribution by warm dust in thermal equilibrium to the dust luminosity of stellar birth clouds
M_{dust}/M_o	Total mass of dust

Many of these features are critical to understanding the star formation rate (SFR) of a galaxy. For instance, large dust luminosity or dust mass values (Cucciati et al. 2012) could lead to an increased SFR, as a larger amount of dust can provide the necessary conditions for more frequent star formation. Dust can shield newly forming stars from harsh radiation, allowing them to accumulate mass and grow. Metallicity, which refers to the abundance of elements heavier than hydrogen and helium, plays a crucial role (Yates et al. 2012) - higher metallicity can enhance cooling processes in molecular clouds, promoting star formation. Furthermore, the star formation history of a galaxy offers valuable context (Boquien et al. 2014), as galaxies with a history of active star formation are likely to continue forming stars at a significant rate. The age of the oldest stars in a galaxy also provides insights (Graus et al. 2019), as older stellar populations can influence the current interstellar medium's dynamics and composition. Values such as the equilibrium temperature of the ISM and the contribution of Polycyclic Aromatic Hydrocarbons (PAHs) could influence the star formation rate (Peeters et al. 2004), as cooler dust will collapse to form stars easier than hotter dust. By examining these features, we gain a comprehensive understanding of the factors driving SFR, allowing us to predict future star formation activity more accurately. This holistic approach highlights the interplay between various galactic properties and underscores the complexity of modeling SFR.

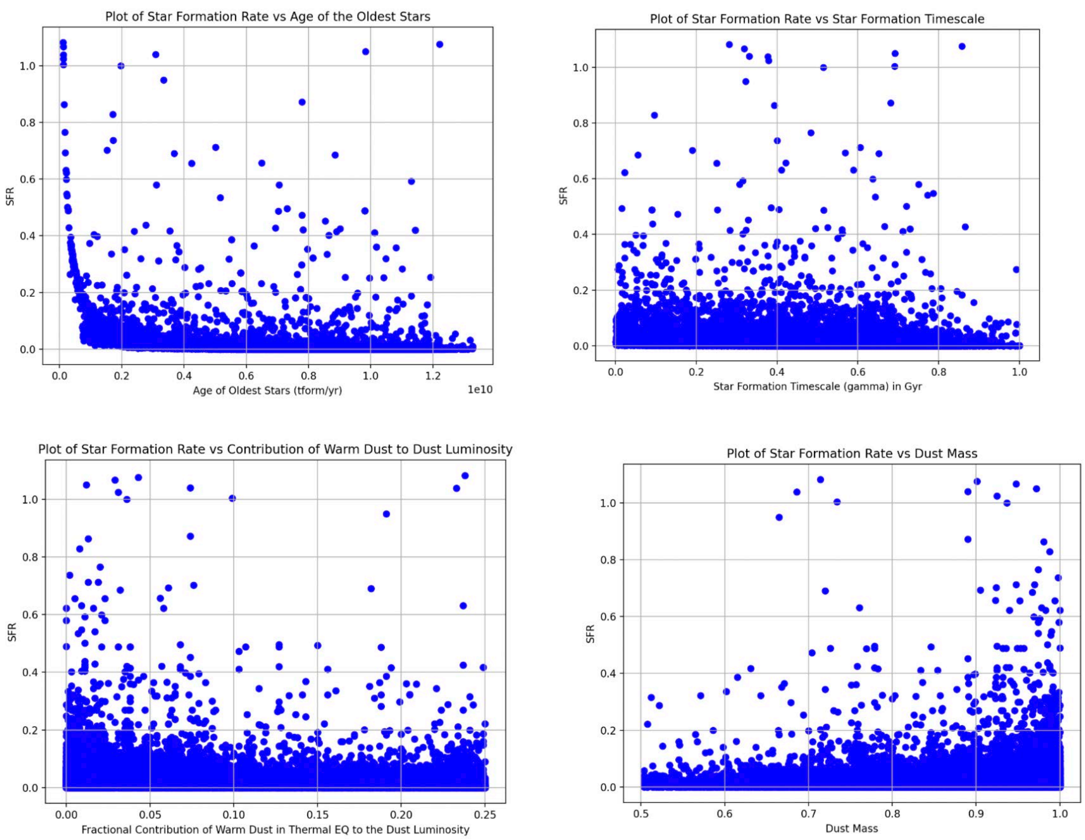

Figure 1. Plots of star formation rate against the age of the oldest stars, star formation timescale, contribution of warm dust to dust luminosity, and dust mass

Methods

For our data, we took the MagPhys output by assigning our X-values to all columns excluding the star formation rate, and our y-values to the star formation rate. We then split our data halfway, 50% being training and 50% being testing, and assigned a random state of 42.

In our research, we leverage linear regression to analyze the relationship between a continuous dependent variable (y) and one or more independent variables (x). This statistical method is expressed through the equation $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \cdots + \beta_n x_n + \epsilon$, where the variables represent the following values as shown in Table 2.

Table 2. Variable Names for Linear Regression

Variable:	Definition:
y	predicted value of our dependent variable (star formation rate)
β_0	Intercept
β_n	Coefficient of the independent variable (indicates the change in y per unit change in x) for each independent variable
ε	Error term (represents the difference between actual and predicted values)

We employ linear regression to uncover the strength and direction of the linear association between the variables. By minimizing ε, we establish the best-fitting straight line that explains the relationship between x and y. This technique allows us to predict y, in our case the star formation rate, for new data points within the modeled range of X, which are our other galactic features.

In addition to standard linear regression, we also utilize Lasso (Least Absolute Shrinkage and Selection Operator) regression (Tibshirani 1996), a variation of linear regression that enhances model performance by incorporating regularization. Lasso regression modifies the linear regression equation by adding a penalty term to the loss function, which is the sum of the absolute values of the coefficients, modeled by the equation $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \cdots + \beta_n x_n + \lambda \sum_{j=1}^{p} |\beta_j| + \epsilon$, where the additional variables represent the following values as shown in Table 3.

Table 3. Additional Variables in Lasso Regression

Variable:	Definition:
λ	Regularization parameter that controls the strength of the penalty
$\lambda \sum_{j=1}^{p} \lvert\beta_j\rvert$	Sum of the absolute values of the coefficients

The penalty term in Lasso regression has several key benefits: it can reduce some coefficients to zero, effectively selecting the most significant features and improving model interpretability by focusing on fewer, more relevant variables. Additionally, by penalizing large coefficients, Lasso regression helps prevent overfitting, especially in high-dimensional data. This approach allows us to better manage data complexity, identify crucial galactic features, and enhance the accuracy and robustness of our star formation rate predictions.

In our research, we use a neural network to study the relationship between the star formation rate (our dependent variable, y) and various galactic features (our independent variables, x). Neural networks are ideal for this task because they can capture complex, non-linear relationships that simpler methods might miss. This allows us to make more accurate predictions and gain better insights into what influences star formation.

A neural network consists of layers of interconnected neurons. Each neuron applies a linear transformation followed by a non-linear activation function to its input. This allows the network to learn patterns in the data. The network is trained by adjusting the weights and biases of the neurons to minimize a loss function, which measures the difference between predicted and actual values. We also use two activation functions in our neural network, which are ReLU (Agarap 2018) and Sigmoid (Sharma et al. 2020).

ReLU (Rectified Linear Unit) introduces non-linearity by outputting zero for negative inputs and the input itself for positive inputs. The function is defined as $ReLU(x) = max(0, x)$, where if x > 0, f(x) = x, and otherwise, f(x) = 0. This property helps the network learn complex patterns and mitigates the vanishing gradient problem, which is common in deep networks. Most of our layers use ReLU to add the necessary non-linearity and improve learning efficiency. Sigmoid normalizes the output to a range between 0 and 1, making it useful for certain layers where output values need to be bounded. For very large positive x values, the function goes to 1, and for very large negative x values, the function goes to 0. For x = 0, the function is at 0.5. The function is defined as $\sigma(x) = \frac{1}{1+e^{-x}}$. By normalizing the outputs, Sigmoid helps control the data's range as it passes through the network, particularly in layers where we want the output to represent probabilities or scaled values.

Our neural network has the following structure, where W_i and b_i are the weights and biases for each layer:

Table 4. Hidden Layers for Neural Network

Layer:	Neurons:	Activation:	Transformation:
Input	16	N/A	N/A
1	186	ReLU	$h_1 = ReLU(W_1 * x + b_1)$
2	186	ReLU	$h_2 = ReLU(W_2 * h_1 + b_2)$
3	93	ReLU	$h_3 = ReLU(W_3 * h_2 + b_3)$
4	93	Sigmoid	$h_4 = Sigmoid(W_4 * h_3 + b_4)$
5	93	ReLU	$h_5 = ReLU(W_5 * h_4 + b_5)$
6	62	ReLU	$h_6 = ReLU(W_6 * h_5 + b_6)$
7	62	Sigmoid	$h_7 = Sigmoid(W_7 * h_6 + b_7)$
8	62	ReLU	$h_8 = ReLU(W_8 * h_7 + b_8)$
Output	1	N/A	$y = W_9 * h_8 + b_9$

To train our neural network, we used the RMSProp (Root Mean Square Propagation) optimizer. RMSProp (Ruder 2016) is an adaptive learning rate optimization algorithm that adjusts the learning rate based on the average of recent squared gradients. This helps the network learn more efficiently and converge faster.

Using this neural network architecture and the RMSProp optimizer provides several benefits. The use of ReLU and Sigmoid functions allows the model to capture complex, non-linear relationships between features and the star formation rate. Multiple layers help the model learn intricate interactions between features, enhancing its ability to understand the data. Additionally, RMSProp helps prevent overfitting by adjusting learning rates dynamically based on observed gradients, leading to more stable and efficient learning. This approach enhances our ability to make accurate predictions of the star formation rate, providing deeper insights into the underlying astrophysical processes.

Mean Squared Error (MSE), R-squared (R^2), and Standard Deviation (SD) are fundamental metrics in statistical analysis. MSE is calculated by averaging the squared differences between observed and predicted values, serving as a measure of prediction accuracy. Lower MSE values indicate a model with better predictive performance. R^2, the coefficient of determination, represents the proportion of variance in the dependent variable that is predictable from the independent variables, with values closer to 1 indicating a better fit. Standard Deviation quantifies the dispersion of data points around the mean, providing insight into data variability.

In the context of determining star formation rates (SFR), these metrics are crucial for model evaluation. MSE helps in assessing how well the model predicts SFR, with lower values indicating more accurate predictions. R^2 indicates how much of the variance in SFR is explained by the model, helping understand the model's explanatory power. Standard Deviation helps in assessing the natural variability in SFR data, distinguishing between model errors and inherent data spread.

Feature importance determines the impact of individual input features on a model's predictions, aiding in model interpretation and refinement. It is useful for understanding data relationships, identifying irrelevant features, and improving model performance.

In linear regression, feature importance is assessed through the absolute values of the model's coefficients. As mentioned earlier, the equation for linear regression is $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \cdots + \beta_n x_n + \epsilon$ where β_i are the coefficients. Higher absolute values of β_i indicate greater importance of the corresponding feature x_i.

In Lasso regression, feature importance is determined by adding an L1 regularization term to the loss function, penalizing the absolute size of the coefficients. The loss function for Lasso regression is $Loss = RSS + \lambda \sum_{j=1}^{p} |\beta_j|$ where RSS is the residual sum of squares and λ is the regularization parameter. This penalty causes some coefficients to shrink to zero, effectively performing feature selection by retaining features with non-zero coefficients as important and setting others to zero.

Results and Discussion

Table 5. Error Values for each model

Model:	Mean Squared Error:	R^2:	Standard Deviation:
Linear Regression	0.00167	0.18221	0.04083
Lasso Regression (alpha = 0.1)	0.00169	0.17076	0.04111
Neural Network	0.000939	0.4808	0.03003

The feature importance of the Linear Regression model showed that the fractional contribution by the hot mid-infrared continuum to the dust luminosity of stellar birth clouds had the largest coefficient of 0.05. Other significant factors in the Linear Regression model included the star formation timescale, with a coefficient of 0.009, and the fractional contribution by warm dust in thermal equilibrium to the dust luminosity of stellar birth clouds, with a coefficient of 0.0013.

Because the Linear Regression model performed better with no regularization, we judged that applying Lasso Regression didn't improve the model's accuracy. However, the feature importance showed valuable results. Because Lasso Regression tends to decrease the importance of many factors, only two factors had a pronounced effect. The largest was dust mass, with a coefficient of about 7×10^{-12}, and the second was the age of the oldest stars, with a coefficient of about 5×10^{-12}. These coefficients compare to an average coefficient of about 7×10^{-13}, showing that the aforementioned features are an order of magnitude more important than the average feature. This makes sense scientifically, as a galaxy with more dust will likely have more material to form stars, increasing the SFR, and a galaxy with more old stars might have more stars in general, decreasing the available gas and dust for forming new stars, and therefore decreasing the SFR.

After compiling our neural network, we observed a mean squared error of 0.000939, representing a decrease of 56% from the mean squared error of 0.00167 observed in the Linear Regression model. More importantly, the linear regression model has an R^2 value of 0.18221, while the neural network has an R^2 value of 0.4808. This shows that the linear regression model has an almost uncorrelated prediction, while the neural network model has a somewhat correlated prediction that is significantly closer to making an accurate prediction. Furthermore, the plots in Figure 1 show that the correlation between different galactic features and star formation rate is nonlinear, suggesting that the neural network picks up on patterns that may not be discovered otherwise. The neural network's stronger correlation shows that more complex models could predict star formation rates better, suggesting that increasing the data size or complexity of the model has potential for accurately predicting star formation rates.

While the neural network, with its superior R^2 value, can predict star formation rate better than the linear regression model, our inability to represent feature importance in neural networks limits our understanding of which specific galactic properties are most influential in determining SFR. Despite these limitations, our results highlight the necessity of using sophisticated models to accurately predict SFR, showing the potential of neural networks to uncover intricate patterns that simpler models may miss. Moving forward, ensuring the accuracy of input data and exploring methods to interpret neural network results, such as explainable AI techniques, will provide deeper insights into the factors driving star formation.

Additionally, our model was trained on various galactic properties derived from the MagPhys model, which analyzes spectral energy distributions to infer these properties. While our neural network showed a significant reduction in mean squared error compared to the linear regression model, we must consider potential limitations. If the MagPhys model's analysis is inaccurate, the machine learning model could learn from incorrect values, leading to misleading conclusions about the relationship between galactic properties and SFR. Thus, the reliability of our predictions depends heavily on the accuracy of the underlying MagPhys model.

Conclusion

We used Linear Regression, Lasso Regression, and a Neural Network to determine the Star Formation Rate from galactic factors such as luminosity, metallicity, the ISM, star formation history, and the masses of gas, dust, and stars. We did this through the use of the Galaxy and Mass Assembly Dataset and the MagPhys model, which derives galactic features from the spectral energy distributions of galaxies and their redshifts. Our models achieved mean squared error accuracies of 0.00167, 0.00169, 0.000939, respectively. We also had a low standard deviation of 0.04083 for linear regression, which is an improvement over previous models. This high accuracy demonstrated that the star formation rate could be accurately predicted by knowing galactic features. It also demonstrated the importance of some features such as star formation timescale, fractional contribution of warm dust at thermal equilibrium to the dust luminosity, the dust luminosity, the dust mass, and the age of the oldest stars.

Exploring methods like those used by Surana et al. 2020. for determining SFR, stellar mass, and dust luminosity is also promising. Their machine-learning techniques could be used to find more galactic features accurately and quickly. This would help study the relationships between these features and SFR in greater detail, leading to more precise models and a better understanding of star formation processes. Their model's increased speed will also allow for faster characterization of more galaxies, allowing for a more representative dataset. Combining an expanded feature set with advanced methods could significantly improve the ability to predict and understand star formation in galaxies.

In future research, adding more galactic features could improve the accuracy of SFR predictions. Factors like magnetic fields (Hocuk et al. 2012), which affect the collapse of molecular clouds, and the galactic environment, including interactions with nearby galaxies, can impact SFR. Radiation pressure (Rosen et al. 2020) from young, massive stars can disperse gas, and cosmic ray flux (Neronov et al. 2017) can change the chemistry of the interstellar medium, both important for star formation. Other, less studied features could be investigated as well, such as the impact of dark matter, black hole activity (Harrison 2017), and galaxy mergers on star formation. Including these factors would help develop a more complete understanding of the mechanisms driving star formation. This approach could lead to new insights into the life cycle of galaxies and the evolution of the universe.

Due to computational limitations, our research also used only 30000 data points from nearly 200000. This could be a limited representation of galaxies as only 15% of the possible data was used. Using more data points, such as 100000 or the full ~200000 galaxies, would serve as a more accurate depiction of star formation rates.

Acknowledgments

I would like to thank the team at Inspirit AI for their support and resources, and I am especially grateful to my coauthor and mentor for their guidance and mentorship.

References

Agarap, A. F. M. (2018). *Deep learning using rectified linear units (ReLU)*. arxiv. https://doi.org/10.48550/arXiv.1803.08375

Bonjean, V., Aghanim, N., Salomé, P., Beelen, A., Douspis, M., & Soubrié, E. (2019, February 8). *Star formation rates and stellar masses from machine learning*. Astronomy & Astrophysics. https://doi.org/10.1051/0004-6361/201833972

Boquien, M., Buat, V., & Perret, V. (2014, November 13). *Impact of star formation history on the measurement of star formation rates*. Astronomy & Astrophysics. https://doi.org/10.1051/0004-6361/201424441

Calzetti, D. (2007, July 3). *Star formation rate determinations*. arXiv.org. https://doi.org/10.48550/arXiv.0707.0467

Colless, M. (1999). *First results from the 2df galaxy redshift survey*. 2dfgrs.net. https://doi.org/10.48550/arXiv.astro-ph/9804079

Cucciati, O., Tresse, L., Ilbert, O., Fèvre, O. L., Garilli, B., Brun, V. L., Cassata, P., Franzetti, P., Maccagni, D., Scodeggio, M., Zucca, E., Zamorani, G., Bardelli, S., Bolzonella, M., Bielby, R. M., McCracken, H. J., Zanichelli, A., & Vergani, D. (2012, February 23). *The star formation rate density and dust attenuation evolution over 12 Gyr with the VVDS surveys*. Astronomy & Astrophysics. https://doi.org/10.1051/0004-6361/201118010

Cunha, E. da, & Charlot, S. (n.d.). *MAGPHYS multi-wavelength analysis of galaxy physical* ... iap.fr. http://www.iap.fr/magphys/ewExternalFiles/readme.pdf

Driver, S. P., Bellstedt, S., Robotham, A. S. G., Baldry, I. K., Davies, L. J., Liske, J., Obreschkow, D., Taylor, E. N., Wright, A. H., Alpaslan, M., Bamford, S. P., Bauer, A. E., Bland-Hawthorn, J., Bilicki, M., Bravo, M., Brough, S., Casura, S., Cluver, M. E., Colless, M., … Wilkins, S. M. (2022, June). *Galaxy and Mass Assembly (GAMA): Data release 4 and the Z < 0.1 total and z < 0.08 morphological galaxy stellar mass functions*. NASA/ADS. https://doi.org/10.1093/mnras/stac472

Eisenstein, D. J., Weinberg, D. H., Agol, E., Aihara, H., Allende Prieto, C., Anderson, S. F., Arns, J. A., Aubourg, É., Bailey, S., Balbinot, E., Barkhouser, R., Beers, T. C., Berlind, A. A., Bickerton, S. J., Bizyaev, D., Blanton, M. R., Bochanski, J. J., Bolton, A. S., Bosman, C. T., … Zhao, B. (2011, September). *SDSS-III: Massive spectroscopic surveys of the distant universe, the milky way, and extra-solar planetary systems*. NASA/ADS. https://doi.org/10.48550/arXiv.1101.1529

Graus, A. S., Bullock, J. S., Fitts, A., Cooper, M. C., Boylan-Kolchin, M., Weisz, D. R., Wetzel, A., Feldmann, R., Faucher-Giguère, C.-A., Quataert, E., Hopkins, P. F., & Keres̆, D. (2019, September 21). *A predicted correlation between age gradient and star formation history in fire dwarf galaxies*. OUP Academic. https://doi.org/10.1093/mnras/stz2649

Hocuk, S., Schleicher, D. R. G., Spaans, M., & Cazaux, S. (2012, September 4). *The impact of magnetic fields on the IMF in star-forming clouds near a supermassive black hole*. Astronomy & Astrophysics. https://doi.org/10.1051/0004-6361/201219628

John, M., Tremonti, C. A., & Kennicutt, R. C. (2006, May 10). *Optical star formation rate indicators*. The Astrophysical Journal. https://doi.org/10.48550/arXiv.astro-ph/0511730

Liske, J., Lemon, D. J., Driver, S. P., Cross, N. J. G., & Couch, W. J. (2003, September). *The Millennium Galaxy Catalogue: 16 <=Bmgc < 24 galaxy counts and the calibration of the local galaxy luminosity function*. NASA/ADS. https://doi.org/10.1046/j.1365-8711.2003.06826.x

McKee, C. F., & Ostriker, E. C. (2007, September 22). *Theory of star formation*. Annual Review of Astronomy and Astrophysics. https://doi.org/10.48550/arXiv.0707.3514

Neronov, A., Malyshev, D., & Semikoz, D. V. (2017, September 29). *Cosmic-Ray Spectrum in the local galaxy*. Astronomy & Astrophysics. https://doi.org/10.1051/0004-6361/201731149

Peeters, E., Spoon, H. W. W., & Tielens4, A. G. G. M. (2004, October 1). *IOPscience*. The Astrophysical https://doi.org/10.1086/423237

Rosen, A. L., Krumholz, M. R., https://orcid.org/0000-0003-4423-0660, A. L. R., & https://orcid.org/0000-0003-3893-854X, M. R. K. (2020, July 23). *IOPscience*. The Astronomical Journal. https://doi.org/10.48550/arXiv.2006.04829

Ruder, S. (2016). *An overview of gradient descent optimization algorithms*. arxiv. https://doi.org/10.48550/arXiv.1609.04747

Sharma, S., Sharma, S., & Athaiya, A. (2020). *Activation functions in neural networks*. ijeast.com. https://ijeast.com/papers/310-316,Tesma412,IJEAST.pdf

Surana, S., Wadadekar, Y., Bait, O., & Bhosale, H. (2020, February 26). *Predicting star formation properties of galaxies using Deep Learning*. OUP Academic. https://doi.org/10.1093/mnras/staa537

Tibshirani, R. (1996). *Regression shrinkage and selection via the lasso*. JSTOR. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Yates, R. M., Kauffmann, G., & Guo, Q. (2012, April 17). *The relation between metallicity, Stellar Mass and star formation in galaxies: An analysis of observational and model data*. OUP Academic. https://doi.org/10.1111/j.1365-2966.2012.20595.x