EXOPLANET DETECTION

USING DEEP LEARNING AND MACHINE LEARNING

SHARON R
B.TECH(ADS)

Agenda

ABSTRACT

The project's primary aim is to harness the capabilities of machine learning and deep learning methodologies in order to meticulously analyze flux data derived from stars and effectively classify them into two categories: exoplanet-stars or non-exoplanet-stars. This classification task hinges on the development of a model that can acquire knowledge from a provided dataset, learning to distinguish stars that are accompanied by exoplanets from those that are not. The project unfolds in several key stages. First and foremost, a comprehensive dataset is gathered, comprising essential information about stars, particularly their variations in light intensity, or flux data. This dataset serves as the fundamental basis for training the machine learning model. Subsequently, machine learning and deep learning models are trained using this dataset, endowed with the ability to detect patterns and features within the flux data that are indicative of exoplanet presence. During the training phase, these models also extract relevant features from the data, such as periodic fluctuations or unique characteristics that differentiate exoplanet-hosting stars from their counterparts.

The models are meticulously schooled to recognize and understand the intricate patterns and anomalies within the data. Once the models are adequately trained, they become proficient at classifying stars, even when presented with data from stars they haven't encountered before. This predictive ability is grounded in the patterns and features they've learned during training. To ensure the accuracy and robustness of the model, validation and testing procedures are conducted, employing separate datasets to assess its performance and make necessary adjustments. In summary, this project represents a cutting-edge approach to unraveling the mysteries of the cosmos, leveraging advanced machine learning techniques to distinguish stars with exoplanets from those without, with profound implications for the field of astronomy and the ongoing quest to uncover exoplanets in the vast expanse of the universe.
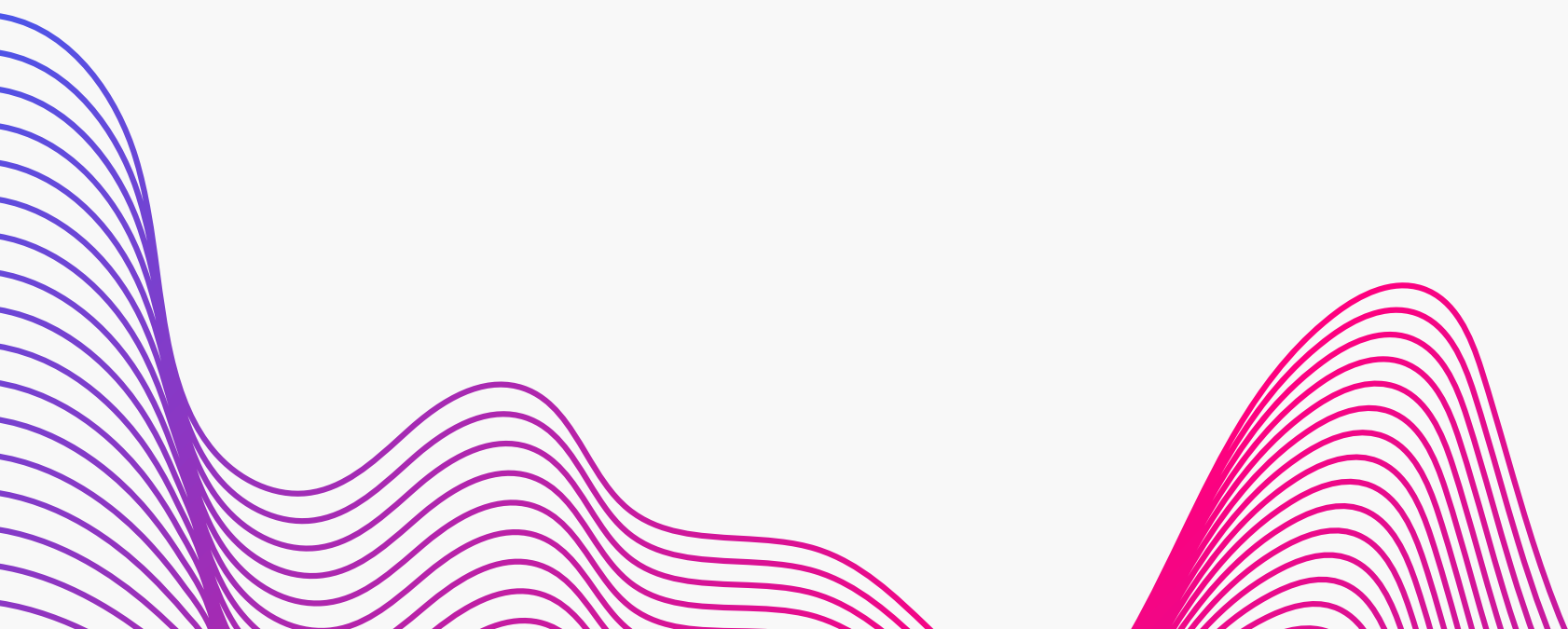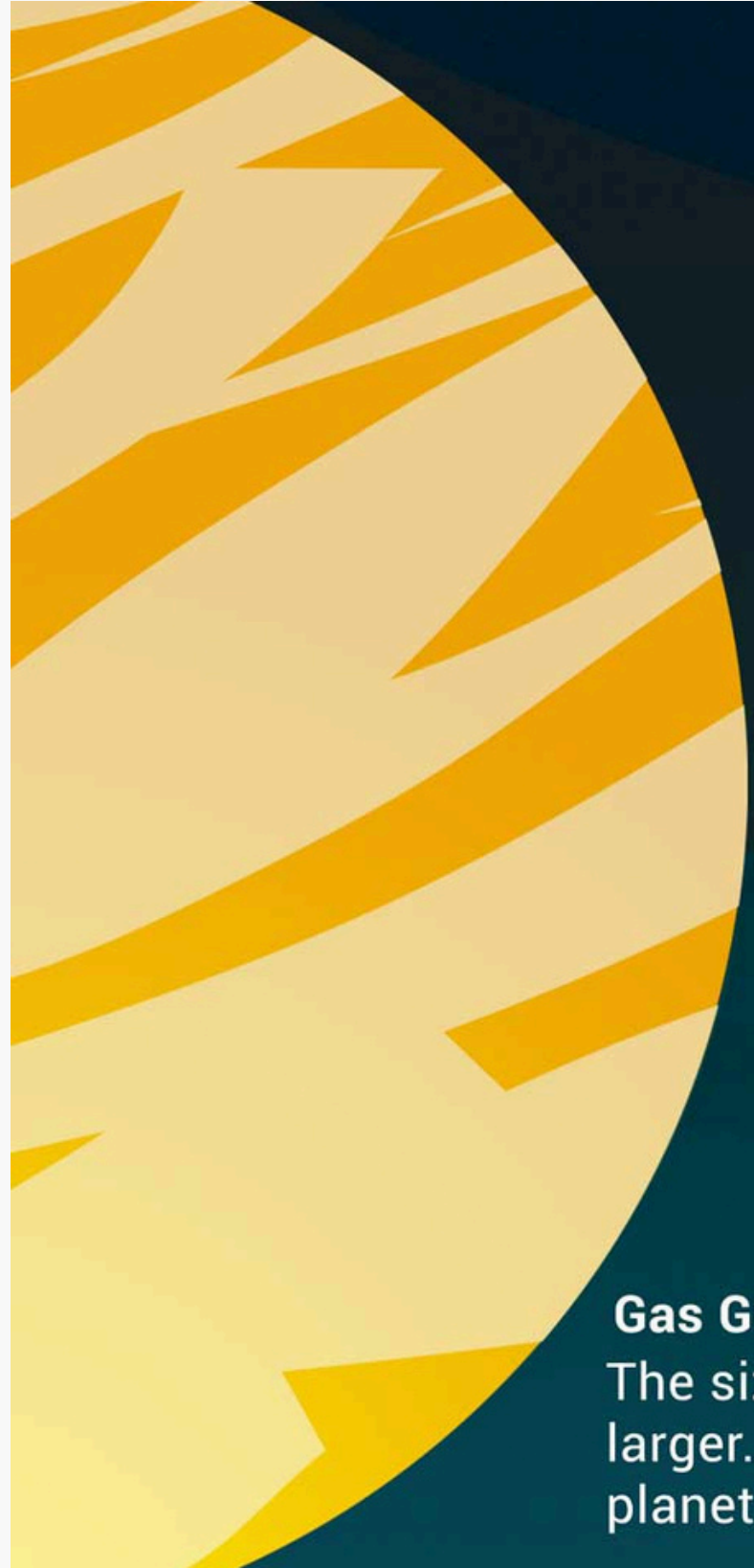
INTRODUCTION

WHAT IS EXO PLANET?

- An exoplanet, or extrasolar planet, is a planet that orbits a star outside of our solar system. These celestial bodies are of great scientific interest as they provide valuable insights into the formation, composition, and diversity of planetary systems beyond our own.

- Exoplanets can vary in size, composition, and orbital characteristics, ranging from gas giants to rocky planets. Their detection is achieved through various indirect methods, such as observing the transit of a planet in front of its host star or measuring the gravitational influence on the star.

- The study of exoplanets plays a crucial role in advancing our understanding of planetary systems and the potential for extraterrestrial life.

Exoplanet Types

Terrestrial
Earth-sized or smaller, mostly made of rock and metal. Some could possess oceans or atmospheres and perhaps other signs of habitability.

Neptune-Like
Similar in size to our own Neptune and Uranus, with hydrogen or helium-dominated atmospheres. "Mini-Neptunes," not found in our solar system, are smaller than Neptune but larger than Earth.

Super-Earth
Typically "terrestrial," or rocky, and more massive than Earth but lighter than Neptune. They might or might not have atmospheres.

Gas Giants
The size of Saturn or Jupiter, or much larger. They include "hot Jupiters"- scorching planets in close orbits around their stars.

LITERATURE SURVEY

The research paper titled "A Convolutional Neural Network (CNN) Based Ensemble Model for Exoplanet Detection" by Priyadarshini, Ishaani, and Puri, Vikram, published in the journal "Earth Science Informatics" in 2021, presents a significant contribution to the field of exoplanet detection.

Exoplanet detection is a critical area of research in astronomy, and traditional methods often struggle to keep pace with the growing volume of astronomical data.

The paper addresses this challenge by introducing a novel ensemble model that leverages the power of Convolutional Neural Networks (CNNs).

CNNs are well-known for their ability to recognize patterns and features in large datasets, and their application to exoplanet detection is a noteworthy approach. What sets this work apart is the ensemble model, which combines multiple CNNs to enhance detection accuracy.

By pooling the strengths of different networks, the authors aim to boost overall detection rates while minimizing false positives. This research paves the way for further developments in the domain of exoplanet detection, offering a promising avenue for improving our understanding of the cosmos through the application of deep learning techniques in astronomy.

What sets this work apart is the ensemble model, which combines multiple CNNs to enhance detection accuracy. By pooling the strengths of different networks, the authors aim to boost overall detection rates while minimizing false positives.

This research paves the way for further developments in the domain of exoplanet detection, offering a promising avenue for improving our understanding of the cosmos through the application of deep learning techniques in astronomy.

EXISTING SYSTEM

- The existing system for exoplanet detection combines traditional astronomical techniques, such as the transit and radial velocity methods, with advanced data analysis methods and machine learning approaches.

- These methods are essential for sifting through the vast volume of astronomical data to identify exoplanets orbiting stars beyond our solar system. The transit method involves monitoring a star's brightness over time, detecting the temporary dimming caused when an exoplanet passes in front of it.

- The radial velocity method, on the other hand, relies on the analysis of periodic shifts in spectral lines caused by the gravitational influence of exoplanets. In recent years, machine learning, including deep learning with neural networks, has gained prominence in exoplanet detection.

- These techniques analyze light curves to identify patterns indicative of exoplanet transits. Additionally, feature extraction, data repositories from missions like Kepler and TESS, follow-up observations, and data analysis software all play crucial roles in this system .

- Researchers use these methods to identify and confirm exoplanets, and the paper "A Convolutional Neural Network (CNN) Based Ensemble Model for Exoplanet Detection" likely introduces an innovative approach or model within this existing system to enhance the accuracy and efficiency of exoplanet detection.

PROPOSED SYSTEM

- The system integrates cutting-edge technologies, including Convolutional Neural Networks (CNNs), exploratory data analysis (EDA), and machine learning(ML), to enhance the accuracy and efficiency of exoplanet detection. In contrast to traditional methods, this system leverages CNNs to extract intricate patterns and features from the flux data of stars. These neural networks are adept at recognizing subtle variations in starlight, such as those caused by exoplanet transits.

- EDA techniques are likely employed to preprocess and visualize the data, facilitating a deeper understanding of the underlying structures and anomalies in the flux data. Furthermore, DL algorithms are incorporated to classify stars as exoplanet-hosting or non-exoplanet-hosting based on the patterns and features identified by the CNNs. The proposed system, through the fusion of CNNs, EDA, and ML, offers a promising path toward more accurate and automated exoplanet detection, potentially reducing the reliance on labor-intensive traditional methods and enabling the processing of increasingly vast datasets in the search for exoplanets.
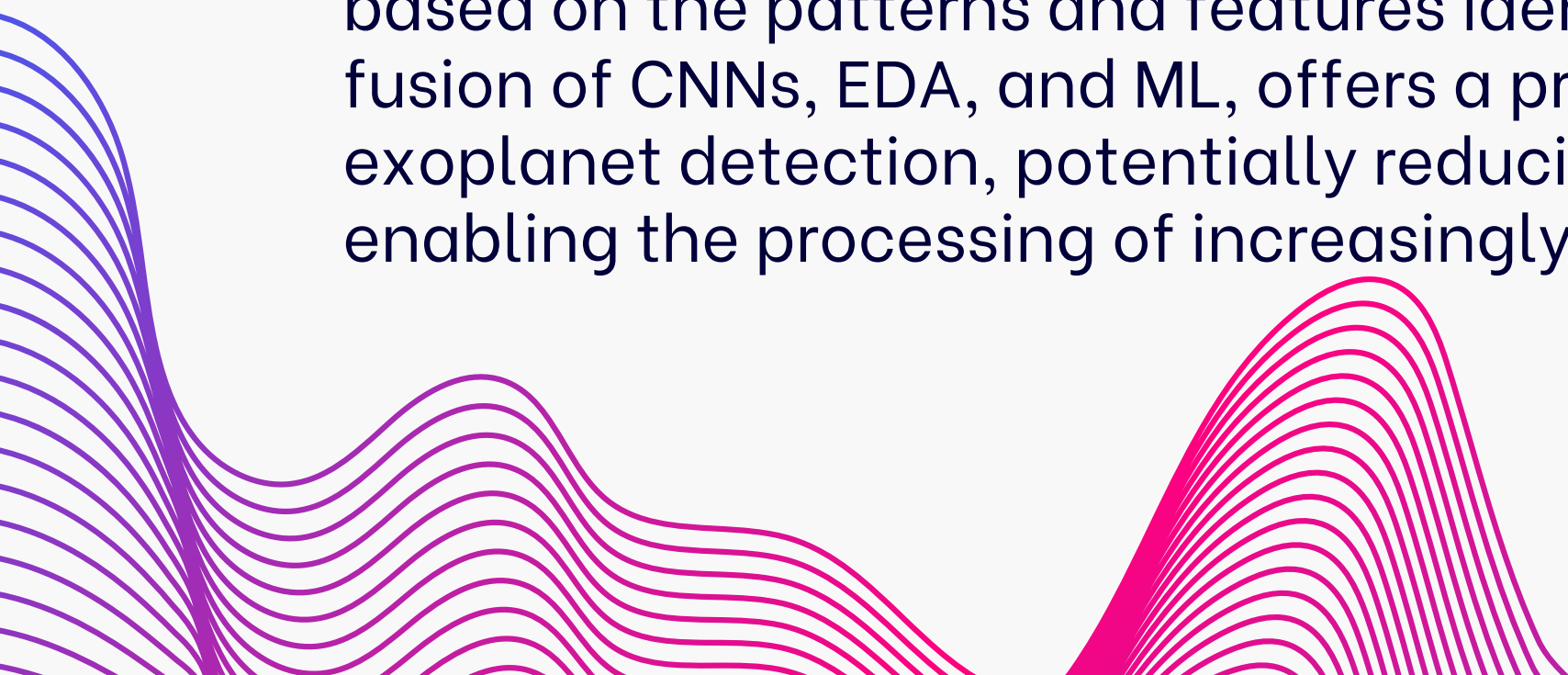
ARCHITECTURE DIAGRAM
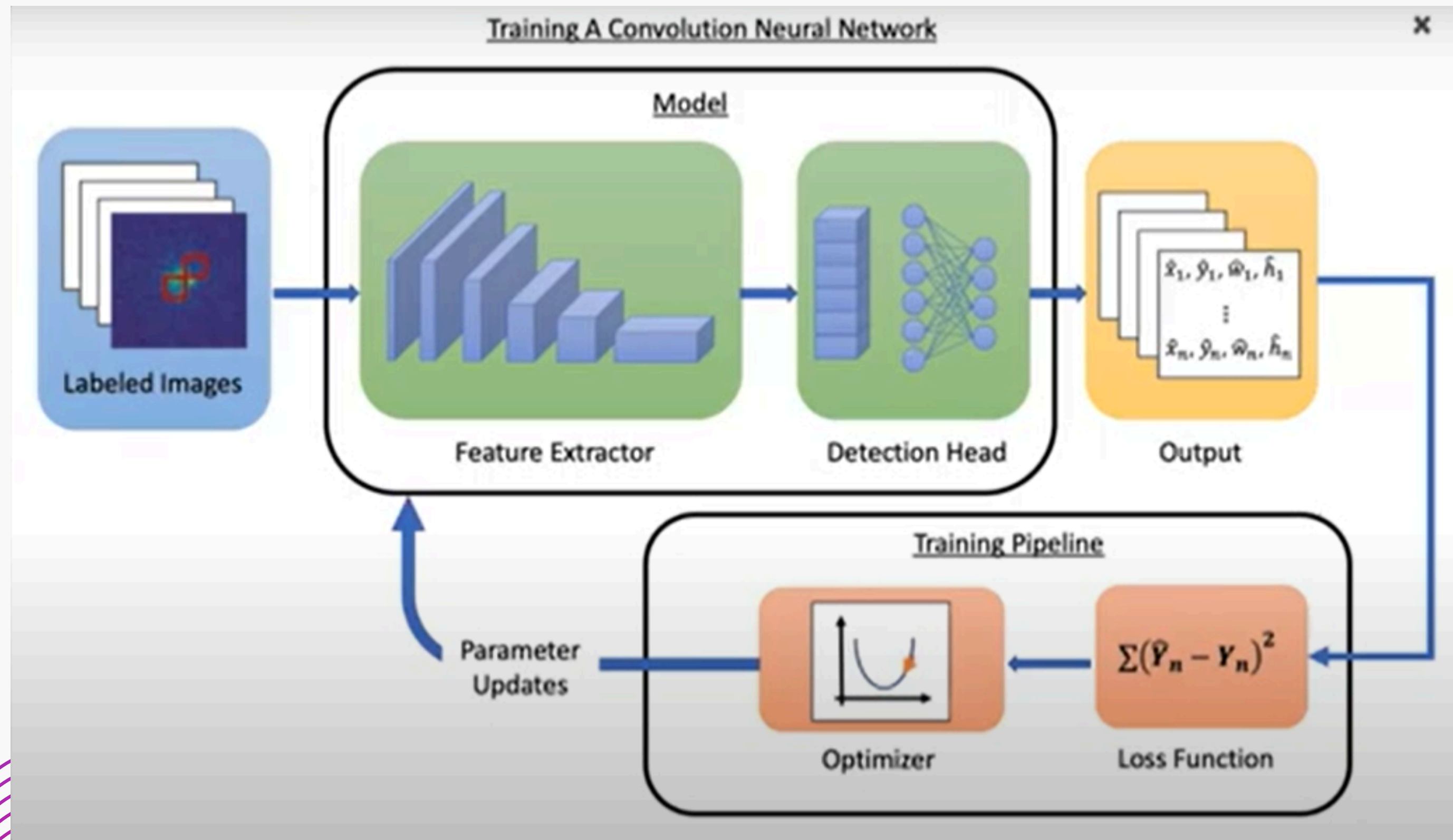
Methods for Detecting Exoplanets

- **Indirect methods:**

  - Transit method: Observing periodic dimming of a star's light as a planet passes in front of it.
  - Radial velocity method: Detecting the wobble of a star caused by the gravitational pull of an orbiting planet.
  - Gravitational microlensing: Measuring the bending of light due to a planet's gravity.
  - Astrometry: Detecting tiny changes in a star's position caused by an orbiting planet.

- **Direct imaging:** Capturing the actual light emitted or reflected by the exoplanet, although challenging due to the brightness of the host star.

Transit Method used for Exoplanet Detection

Flux is a crucial parameter used in the detection and characterization of exoplanets. Flux is a measure of the number of electric or magnetic field lines passing through a surface in a given amount time. By monitoring the flux, which represents the light intensity emitted by a star, astronomers can identify subtle changes that indicate the presence of an exoplanet. The transit method relies on observing periodic dips in flux as an exoplanet passes in front of its host star, causing a temporary decrease in the observed light. Additionally, the radial velocity method measures the small shifts in spectral lines caused by the gravitational tug of an exoplanet, resulting in periodic variations in flux. Analyzing these flux variations provides valuable information about the presence, size, and orbital characteristics of exoplanets.
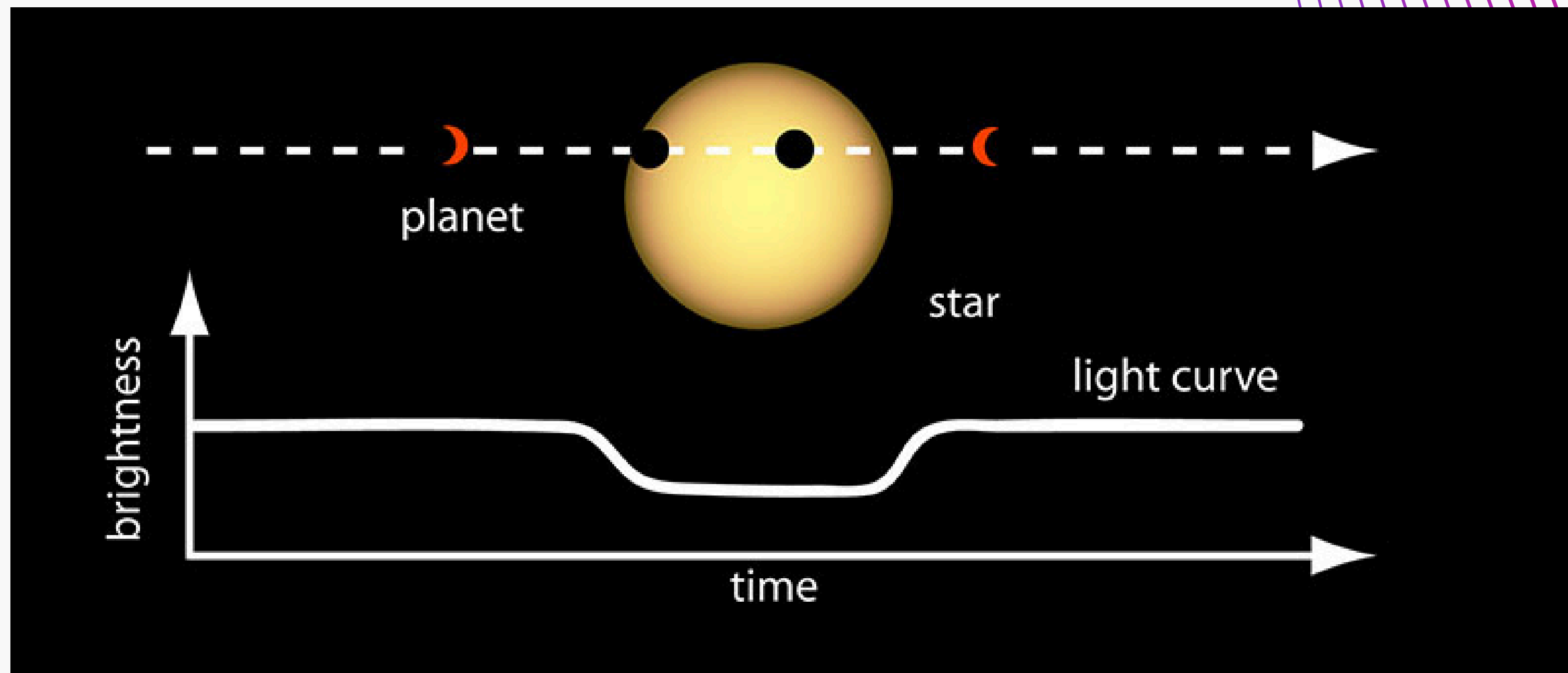
DATASET DESCRIPTION

The dataset for the following project was collected by the NASA Kepler space telescope using the Transit method. By closely observing a star over extended periods, ranging from months to years, scientists can detect regular variations in the light intensity. These variations, known as "dimming," serve as evidence of the presence of an orbiting body around the star.

Such stars exhibiting dimming can be considered potential exoplanet candidates. However, further study and investigation are required to confirm the existence of exoplanets. For example, employing satellites that capture light at different wavelengths can provide additional data to solidify the belief that a candidate system indeed harbors exoplanets.

The dataset provided is divided into Training and Testing data. The data describe the change in flux (light intensity) of several thousand stars. Each star has a binary label of 2 or 1. 2 indicated that that the star is confirmed to have at least one exoplanet in orbit; some observations are in fact multi-planet systems.

- **Trainset:**

    - 5087 rows or observations.
    - 37 confirmed exoplanet-stars and 5050 non-exoplanet-stars.
- **Testset:**

    - 570 rows or observations.
    - 5 confirmed exoplanet-stars and 565 non-exoplanet-stars.

CONCLUSION

- Overall, this project has shed light on the potential of machine learning in exoplanet detection and prediction. It has demonstrated the effectiveness of CNN models and the significance of selecting appropriate algorithms for specific tasks.

- While there is room for improvement in the performance of the machine learning models, this project serves as a valuable foundation for further research in the field of exoplanet exploration and offers insights into the application of machine learning techniques in the domain.

- The project outcomes can guide future research efforts and inspire the development of more accurate and efficient models for exoplanet detection, ultimately advancing our knowledge of the universe and the existence of habitable planets beyond our solar system.