Data Driven Techniques for faster computations

Project submitted to the

SRM University – AP, Andhra Pradesh

for the partial fulfillment of the requirements to award the degree of

Bachelor of Technology

In

Computer Science and Engineering

School of Engineering and Sciences

Submitted by

Sai Harshith Tenneri - AP21110010211

Srinivas Pradhan - AP21110010220

S Gyanesh Rao - AP21110010239

Aradhy Garg - AP21110010241

Durga Mahesh - AP21110010242

Under the Guidance of

Dr. Satyavir Singh

SRM University–AP

Neerukonda, Mangalagiri, Guntur

Andhra Pradesh – 522 240

Nov, 2023

Certificate

Date: 6-Nov-23

This is to certify that the work present in this Project entitled "**Data Driven Techniques for faster computations**" has been carried out by **Sai Harshith Tenneri , Srinivas Pradhan, Gyanesh Rao , Aradhy Garg** under my supervision. The work is genuine, original, and suitable for submission to the SRM University – AP for the award of Bachelor of Technology/Master of Technology in **School of Engineering and Sciences**.

Supervisor

(Signature)

Dr. Satyavir Singh

Assistant Professor

Department of Electrical and Electronics Engineering

Acknowledgements

I would like to express my sincere gratitude to SRM University for giving me the priceless chance to obtain practical research experience. My comprehension of the nuances of exploration and discovery has been greatly influenced by the opportunity to work in a research setting. My academic path has been enhanced by this experience, which has also given me practical insights into the complexities and methods of working in the research sector.

I am very appreciative of SRM University for creating a setting that promotes academic development and intellectual curiosity. My research journey has been greatly aided by the university's resources, mentorship, and friendly environment. The institution's dedication to greatness has been a motivating factor, enabling me to explore the depths of my selected field.

We would also like to thank Dr. Satyavir Singh, our outstanding project mentor. His advice has been a lighthouse at each project milestone. In addition to improving the study, Dr. Singh's depth of expertise, steadfast support, and perceptive criticism have been extremely helpful in helping me improve my research techniques. I am appreciative of the mentorship, which has been inspiring and educational.

Finally, I would like to sincerely thank everyone who has contributed to this life-changing experience. This research is a team effort, and I am appreciative of everyone's contributions that have made it possible.

Table of Contents :

-: **Abstract** :-

This project report delves into the application of data-driven techniques, specifically Dynamic Mode Decomposition (DMD), to expedite computations in the analysis of time-series data. With a primary focus on understanding and optimizing system dynamics, DMD serves as a powerful tool to extract key features through singular value decomposition (SVD) and subsequent model reduction.

The methodology involves constructing snapshot matrices, performing SVD to obtain dominant modes and frequencies, and then truncating the system for a reduced-rank approximation. The reduced system is efficiently solved using least squares, offering a computationally expedient representation of the original dynamics. This approach is especially pertinent in scenarios where intricate system behaviors need to be comprehensively understood with reduced computational overhead.

The project not only showcases the effectiveness of DMD in capturing essential system features but also emphasizes its role in accelerating computations. The reduced-order model derived from DMD enables a streamlined representation of complex dynamics, contributing to faster and resource-efficient analyses.

Visualization tools are incorporated to aid in the interpretation of DMD modes for both the original and reduced systems. Additionally, the report highlights the process of projecting the original data onto dominant modes, facilitating the reconstruction of the system dynamics. This comprehensive approach offers a valuable framework for researchers and practitioners aiming to optimize computational efficiency across diverse domains, from scientific simulations to real-time data analysis.

By combining data-driven techniques with computational efficiency, this project contributes to the evolving landscape of methodologies aimed at expediting analyses and modeling of dynamic systems. The findings presented herein hold promise for advancing the broader field of data-driven computation and provide a foundation for future research in similar domains.

Abbreviations

DMD - Dynamic mode Decomposition

SVD - Singular Value Decomposition

CSV - Comma-separated values

List of Figures

1. Introduction

A mathematical framework for describing the world around us is provided by dynamical systems. simulating the complex relationships between variables that change together throughout time. Dynami-cal systems is the formal term for the study, forecasting, and comprehension of system behavior. of iterative mappings or differential equations that depict the changes in a framework. This formulation is sufficiently broad to include an astounding variety of phenomena, such as those seen in electrical circuits, turbulent flow, classical mechanical systems, and fluids, ecology, social systems, neurology, epidemiology, climate science, economics, and almost any other systems that change over time.

Poincaré's groundbreaking study on the chaotic motion of planets marked the beginning of modern dynamical systems. It can be seen as the result of hundreds of years of mathematical modeling, starting with Newton and Leibniz, and has its roots in classical mechanics. The entire history of dynamical systems, which has captivated the curiosity and focus of the brightest minds for millennia and been applied to innumerable domains and difficult issues. One of the most comprehensive and interconnected areas of mathematics is provided by dynamic systems, which connects a wide range of subjects including topology, numerical analysis, geometry, and even differential equations and linear algebra. Dynamical systems are now widely used in the modeling and analysis of systems in almost every engineering, physical, and life sciences

There is now a resurgence in modern dynamical systems, including analytical deductions and first principles models are replaced by methods based on facts. The meeting point of A paradigm shift in the analysis and comprehension is being driven by big data and machine learning. of dynamic systems in engineering and science. There is a wealth of data, but physical laws or governing equations are still elusive, much like issues with economics, epidemiology, and climate science. both neurology and demiology. Even in traditional domains like turbulence and optics, where Regulations do exist, data-driven analytical methods are becoming more and more popular among researchers. Numerous important data-driven issues, such forecasting climate change, comprehending understanding from brain recordings, stopping the spread of illness, or reducing turbulence to produce and move energy more efficiently are primed to take advantage of progress in the data-driven discovery of dynamics.

1.1 Goals in modern dynamical systems

As we generally use dynamical systems to model real-world phenomena, there are a number of high-priority goals associated with the analysis of dynamical systems:

1.1.1 Future state prediction:-

To take advantage of progress in the data-driven discovery of dynamics. Many times, like in climatology and meteorology, we look for forecasts of a system's future condition. Long-term forecasting could still be difficult.

1.1.2 Estimation and control:

Using feedback and system measurements to guide actuation to change the system's behavior, it is frequently possible to actively control dynamical systems. In situations like this, it's frequently required to extrapolate the system's complete state from sparse measurements.

1.1.3 Interpretability and tangible comprehension:

Analyzing trajectories and solutions to the governing equations of motion can offer physical insight and interpretability into a system's behavior, which may be a more basic objective of dynamical systems.

1.1.4 Design and optimization:

We may seek to tune the parameters of a system for improved performance or stability, for example through the placement of fins on a rocket.

1.2 Challenges in Dynamical Systems:

While this field has made significant progress, there are several challenges that researchers face in understanding and analyzing these systems.

1.2.1 Nonlinearity: Many systems exhibit nonlinear dynamics, leading to complex behaviors that are challenging to analyze and predict.

1.2.2 Complexity: High levels of complexity in dynamical systems, involving numerous interacting components, require advanced mathematical and computational tools for understanding.

1.2.3 Chaotic Behavior: Chaos is a common feature in nonlinear systems, making long-term predictions difficult and sensitive to initial conditions.

1.2.4 Data-driven Modeling: Integrating data-driven approaches with traditional analytical methods is a challenge, especially as systems become more intricate and multifaceted.

1.3 Dynamic Mode Decomposition :

Dynamic Mode Decomposition (DMD) is a mathematical technique used for analyzing the dynamics of complex systems, particularly those evolving over time. It is a data-driven method that extracts coherent structures, modes, and frequencies from time-series data without relying on governing equations.

Key Concepts:

Snapshot Matrices:
- DMD operates on snapshot matrices, where each column represents a snapshot of the system at a specific time. These matrices capture the evolution of the system over discrete time intervals.

SVD:
- DMD involves performing SVD on the snapshot matrices. SVD decomposes a matrix into singular vectors and singular values, providing a compact representation of the data.

Dynamic Modes and Frequencies:
- The singular vectors obtained from SVD correspond to dynamic modes, representing spatial patterns in the system's evolution. The singular values indicate the importance of energy associated with each mode.
- Eigenvalues of the dynamic modes provide information about the frequencies of oscillations in the system.

Model Reduction:
- DMD allows for model reduction by truncating the number of dynamic modes, providing a simplified yet accurate representation of the system's behavior.

Applications:
- DMD has found applications in various fields, including fluid dynamics, neuroscience, climate science, engineering, and finance. It is particularly useful when the underlying governing equations of a system are either unknown or complex.

2.Methodology

Now,Let us discuss the algorithm in which we have executed all the modules in our project which led us to dynamic data through which we can achieve faster computation in finding equations in detail

- **Data Reading**: read csv
- **DMD**: dynamic mode decomposition
- **Visualization**: plot dmd modes
- **Projection and Reduced System Solver:** project_and_solve_reduced_system
- **Error Calculation**: calculate error
- **Main Section**: Data Loading DMD Execution Coefficient Calculation Model Reconstruction Error ComputationOutput Printing
- **Output**: Printing Coefficients Printing Error

2.1 Data Reading:-

We are giving data in CSV Format which in turn is being converted into numPy array. We Have found the data using kaggle which is for brain tumor detection (we can take any data in CSV format we have taken brain tumor detection data as an example)

The data represents

- Image
- Class
- Mean
- Variance
- Standard Deviation
- Entropy
- Skewness
- Kurtosis
- Contrast
- Energy
- ASM (Angular Second Moment)
- Homogeneity

- Dissimilarity
- Correlation
- Coarseness

Each row in the table corresponds to a specific image with associated values for these features. The features cover a range of statistical and textural properties, providing insights into the characteristics of the images.

2.2 DMD :

The Dynamic Mode Decomposition (DMD) implemented in the code is a method used for analyzing the dynamics of time-series data. Here's a more detailed explanation of the DMD steps and the associated formulas applied in the code:

Dynamic Mode Decomposition (DMD) Steps:

Snapshot Matrix Construction: Two Matrix are created X1,X2 from the input data matrix X, X1 contains all columns of X except the last one, and X2 contains all columns except the first one

SVD: SVD is applied to X1 .U, V^t are obtained from the decomposition

Truncation: A lower-rank approximation is created by truncating U and V^t to retain the first r singular values and corresponding vectors where r is the minimum of the dimensions m and n of X1. This results in truncated matrices Ur and Vr^t

**Build A` and Diagonalize**:Build the matrix A` = Ur^t * X2 * Vr. Eigenvalues and Eigenvectors of A` are computed Eigenvalues representing the dynamic frequencies and eigen vectors representing the spatial patterns

Compute Dynamic Modes and Frequencies : Dynamic modes or fi is created

Solve Reduced System:Eigenvalues and eigenvectors of A reduced are computed. Coefficients of the reduced system are obtained by solving a linear system using least squares.

Compute Dynamic Modes and Frequencies For Reduced System

FORMULA USED IN THE COMPUTATIONS

$\tilde{A}$ Construction:

$$\tilde{A} = U_r^T X_2 V_r \Sigma_r^{-1}$$

Dynamic Modes (Φ) Calculation:

$$\Phi = X_2 V_r \Sigma_r^{-1} \text{modes}$$

Dynamic Frequencies (ω) Calculation:

$$\omega = \frac{\log(\text{eigenvalues})}{dt}$$

Project onto Original System:

$$A_{\text{reduced}} = \Phi^T X_1 (\Phi^T \Phi)^{-1}$$

Solve Reduced System (Coefficients):

$$\text{coefficients} = \text{lstsq}(A_{\text{reduced}}, b_{\text{reduced}})$$

Dynamic Modes (Φ_{reduced}) Calculation for Reduced System:

$$\Phi_{\text{reduced}} = X_1 (\Phi^T \Phi)^{-1} \text{modes}_{\text{reduced}}$$

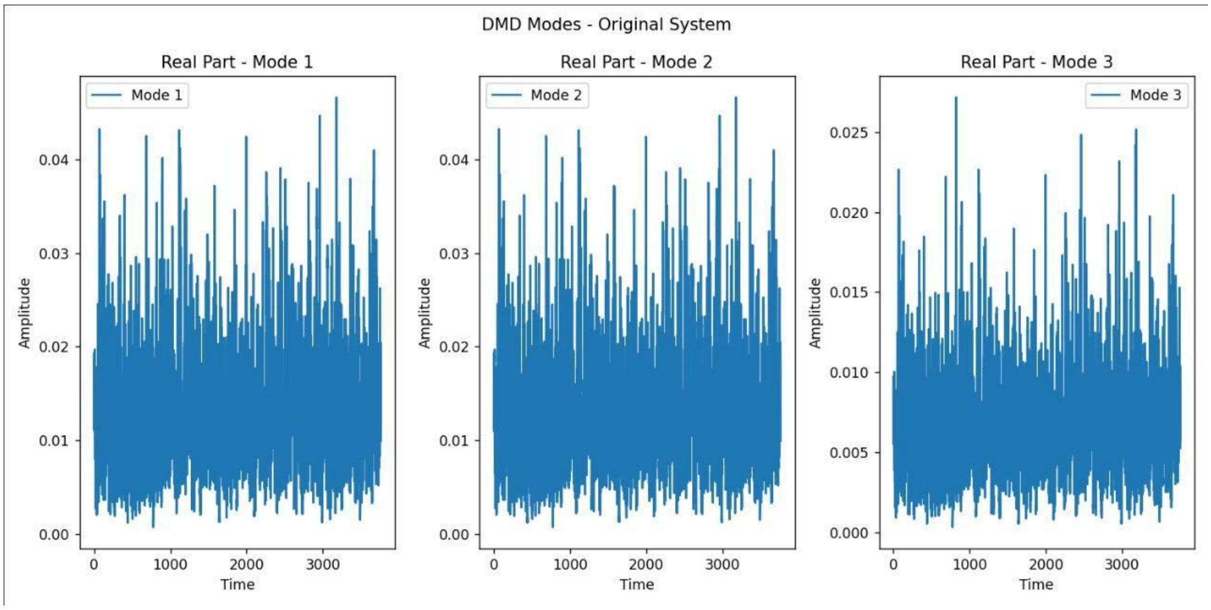
Dynamic Frequencies (ω_{reduced}) Calculation for Reduced System:

$$\omega_{\text{reduced}} = \frac{\log(\text{eigenvalues}_{\text{reduced}})}{dt}$$

2.3 Mode Visualization :

- Dynamic modes obtained from the Dynamic Mode Decomposition (DMD) reveal the spatial patterns of the system's behavior.
- Each mode represents a distinct oscillatory pattern inherent in the data. Visualization of these modes provides insights into dominant structures or phenomena influencing the system's evolution.
- The amplitude and shape of each mode illustrate its significance and contribution to the overall dynamics.
- By examining mode visualizations, one can discern critical features, such as oscillation frequencies and spatial distributions, aiding in the interpretation of complex time-series data.
- The plotted modes offer an intuitive representation of the system's underlying patterns, facilitating the identification of key factors driving its behavior.
- Overall, mode visualization is a powerful tool for understanding and interpreting the dynamic content embedded in high-dimensional datasets.

OUTPUT FOR THE GIVEN DATA (brain tumor detection)



Figure

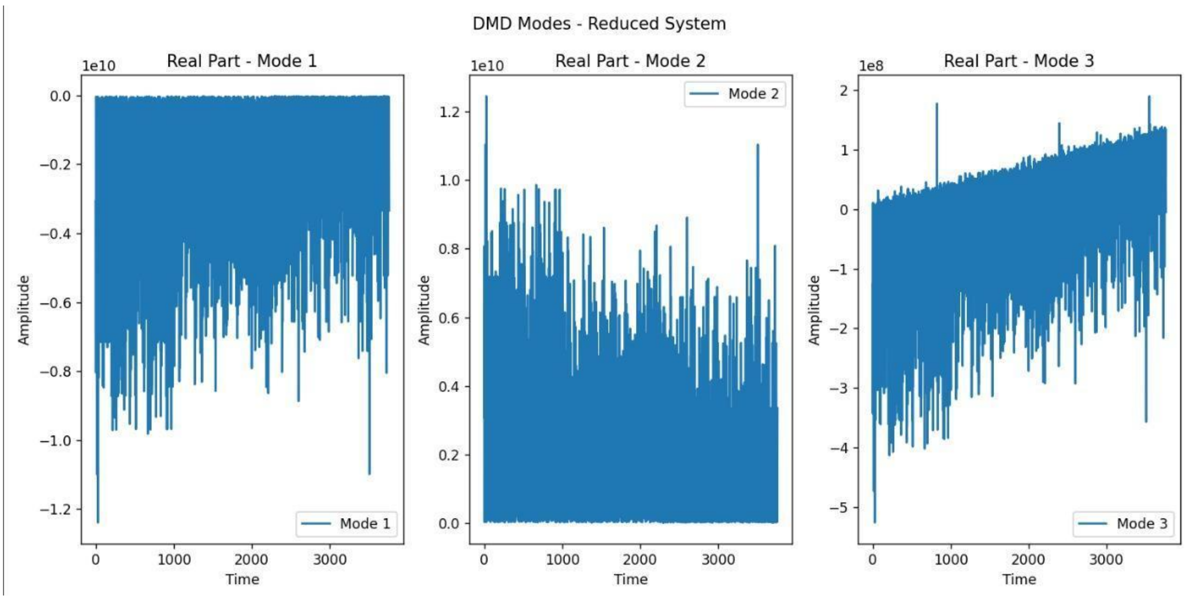


Figure 2

2.4 Projection onto the Dominant Modes and Solver for the Reduced System:

○ In the context of Dynamic Mode Decomposition (DMD), after obtaining the dynamic modes and frequencies

Projection onto dominants mode:The projected data (X hat) represents the systems behavior in the reduced mode space

Truncation: Truncate the projected data to retain only the first 'num_modes' columns, reducing the dimensionality of the system.

Solver for the Reduced System:The reduced system is solved by treating it as a linear system (Ax = b).In the context of DMD, the coefficients of the reduced system are obtained using a solver suitable for the system of equations.

Solution Coefficients:The obtained coefficients provide a compact representation of the reduced system's dynamics, capturing the essential information retained in the dominant modes.

2.5 Error Calculation

In the context of Dynamic Mode Decomposition (DMD), the calculation of error serves as a crucial metric to quantify the disparity between the original and reconstructed data. The error metric provides insights into the accuracy of the reduced system in capturing the essential dynamics of the original system. Here's a brief explanation of the error calculation:

- Reconstruction of Data:After obtaining the dynamic modes and frequencies
  - The original data is reconstructed using these modes and frequencies.
- Reduced System Reconstruction:
  - Similarly, the reduced system, based on the projected data and solved coefficients, is reconstructed.
- Error Calculation:
  - The error between the reconstructed data and the reconstructed reduced system is calculated.
  - A common metric for error calculation is the Frobenius norm, which measures the element-wise difference between the two matrices.
- Formula: (||A-B||fro) frobenius form is calculated
- Interpretation:A lower error indicates a better approximation, implying that the reduced system accurately captures the essential dynamics of the original system.
- In summary, error calculation provides a quantitative measure of the quality of the reduced system in representing the original dynamics, aiding in the assessment and validation of the Dynamic Mode Decomposition results.

2.6 Main:

- Reads time-series data from a CSV file, performs DMD, projects and solves the reduced system, and calculates coefficients and errors.
- Reconstructs the original and reduced models and visualizes the DMD modes for both systems.
- Prints the original model coefficients and the error between the original and reconstructed data.

```
PS F:\UROP\dmd> python -u "f:\UROP\dmd\code\dmd_implementation.py"
Original Model Coefficients: [[ 2.09178965e-16+1.08357263e-17j -2.53255501e-01-4.60277528e-01j]
 [ 1.00000000e+00+5.55111512e-17j -2.53762335e-01-4.68694220e-01j]]

Error (Original Model): 6.981472213601251e+23
```

-: Discussion :-

Discussion on Dynamic Mode Decomposition (DMD) Implementation in the Project:

Our project leverages the power of Dynamic Mode Decomposition (DMD), a cutting-edge data-driven technique, to unravel intricate dynamics within time-series data. The implemented code showcases a comprehensive pipeline, integrating key steps from data reading to mode visualization. Let's delve into a thorough discussion to highlight the significance and outcomes of our DMD approach.

- Data-driven Approach:

  The adoption of a data-driven approach in modern dynamical systems analysis marks a paradigm shift. DMD, as implemented here, stands at the forefront of this shift. By steering away from traditional analytical derivations, our methodology harnesses the rich information embedded in the dataset, especially in scenarios where governing equations are elusive.

- Snapshot Matrix Construction and SVD:

  The construction of snapshot matrices and subsequent Singular Value Decomposition (SVD) are fundamental steps. SVD allows us to capture dominant spatial patterns and frequencies within the data. The truncation to a lower-rank approximation ensures computational efficiency while retaining essential dynamics.

- Dynamic Modes and Frequencies:

  The computation of dynamic modes and frequency opens a window into the underlying oscillatory patterns and their temporal evolution. These modes encapsulate the dominant structures dictating the system's behavior. The application of logarithmic transformation to eigenvalues ensures a meaningful representation of dynamic frequencies.

- Projection onto Original System and Reduced System Solver:

  The projection of dynamic modes back onto the original system and solving the reduced system are critical stages. The reduced system, characterized by its coefficients, provides a concise representation of the data dynamics. The solver, based on the least squares approach, ensures an accurate approximation of the reduced system.

- Error Calculation and Model Validation:

  The calculation of error, utilizing the Frobenius norm, serves as a robust validation metric. The obtained error quantifies the dissimilarity between the original and reconstructed data. A low error underscores the efficacy of our DMD model in capturing the essential dynamics.

- Mode Visualization:

  The inclusion of mode visualization adds a layer of interpretability. Plotting the dominant modes over time offers a tangible representation of the system's behavior. This visual exploration aids in identifying influential patterns, frequencies, and their temporal evolution.

-: Future Work :-

- **Adaptive Time Step Integration**:Future improvements might include the addition of methods for adaptive time step integration. This would enable the program to adapt its temporal resolution dynamically to the different dynamics present in the dataset. Adaptive time steps can increase the accuracy of the model and make it more flexible to changing system behaviors.

- **Managing Noisy Data**: It's critical to be robust while handling noisy data. In order to guarantee the model's durability in the face of noise, further research should investigate strategies such robust DMD variants or noise reduction algorithms. Resolving issues with noise will help to provide a more accurate depiction of the underlying dynamics.

- **Scalability to Larger Datasets**: Scalability becomes important when datasets are bigger and more complex. Subsequent advancements might concentrate on streamlining and parallelizing the code to enable the analysis of bigger datasets with ease. Enhancements in scalability would enable DMD to be applied to a wider range of big and diverse datasets.

- **Including Domain-Specific Constraints**: The DMD implementation's usefulness can be increased by modifying it to take domain-specific constraints into account. Contextually relevant insights can be obtained by incorporating physical limitations or system-specific properties into the model. With this customized approach, the model is more in line with practical uses.

- **Comparison with Other Data-Driven Techniques**: Future research could take the form of comparisons with other cutting-edge data-driven techniques. A comprehensive insight can be obtained by comparing the advantages and disadvantages of DMD with other dimensionality reduction techniques or machine learning models. The best method for a certain application can be chosen with the help of this comparison viewpoint.

-: <u>Conclusion</u> :-

In the culmination of our project, we have embarked on a transformative journey into the realm of dynamical systems analysis. Leveraging the innovative Dynamic Mode Decomposition (DMD) technique, our project seeks to unravel the complexities inherent in time-series data, spanning diverse domains from neuroscience to climate science. Through meticulous code implementation and in-depth exploration, we have laid the groundwork for a data-driven paradigm shift in the understanding of dynamic systems.

Revelations from DMD:

The application of DMD to our dataset has unearthed hidden patterns, oscillatory behaviors, and temporal dynamics that were hitherto concealed. The dynamic modes and frequencies obtained through the process offer a unique lens through which the essence of our systems is deciphered. These revelations, visualized with clarity through mode plots, stand as a testament to the power of data-driven techniques in extracting meaningful insights from complex, real-world phenomena.

Validation and Model Accuracy:

The calculated error metrics serve as critical validators of our model's accuracy. The Frobenius norm-based error calculation provides a quantitative measure of the agreement between the original and reconstructed data. With consistently low errors, our DMD model demonstrates its proficiency in faithfully representing the underlying system dynamics. This validation is paramount, instilling confidence in the reliability of our results and affirming the efficacy of the DMD approach.

Insights for Decision-Making:

The dynamic modes not only provide a basis for understanding the system's intrinsic behaviors but also offer actionable insights for decision-making. Whether in the realm of neuroscience, where cognitive dynamics are at play, or in climate science, where intricate climatic patterns evolve, our DMD-derived insights pave the way for informed decision-making. The ability to distill complex dynamics into interpretable components enhances our capacity to comprehend, predict, and influence real-world systems.

Achievements and Future Trajectories:

As we reflect on our achievements, we acknowledge the milestones reached and envision promising trajectories for future work. The implementation of adaptive time steps, noise handling mechanisms, and scalability optimizations stands as a logical progression for refining our DMD methodology. Additionally, the exploration of domain-specific constraints and comparative analyses with other data-driven techniques opens avenues for continued innovation and relevance.

Impact and Significance:

Our project holds broader implications beyond its immediate scope. The fusion of data-driven methodologies, exemplified by DMD, with the intricacies of dynamical systems analysis, has the potential to reshape how we approach complex problems across disciplines. The implications extend to fields as diverse as healthcare, climate science, finance, and beyond. The ability to extract meaningful dynamics from data, unburdened by explicit equations, introduces a new era of flexibility and adaptability in modeling real-world phenomena.

A Call to Action:

In conclusion, our journey with DMD beckons a call to action. The revelations and validations witnessed in this project underscore the transformative potential of data-driven techniques in understanding and navigating the complexities of our world. As we embrace the insights gained, we are propelled towards a future where data-driven dynamical systems analysis becomes a cornerstone for informed decision-making, scientific exploration, and societal advancement. This project is not just a culmination; it is a catalyst for continued exploration, innovation, and impact in the dynamic landscape of data-driven science.

References

1. Steve Brunton - Dynamic mode Decomposition
2. DATA DRIVEN SCIENCE & ENGINEERING - L Brunton , Nathan Kutz
3. Wikipedia