Venkateswar Reddy M.
CTO, Brillium Technologies
"...dare to dream; care to win..."

Date: **10-Sep-24**

Applied AI & ML Industry Projects Lab (AAM-IPL)
Transforming Learning into Industry Solutions

Week-1 AAM-IPL Project – Ames Housing

B.Tech – CSE(AIML)
V Semester - ML and AIUP, Aug-Oct 2024
Department of Computer Science Engineering – AI and ML (CSM)
G.Pulla Reddy Engineering College (Autonomous), Kurnool, AP

Algorithm of Application: PCA for Dimensionality Reduction

Project Statement

You are given the data set of Ames Assessor's Office used in computing assessed values for individual residential properties sold in Ames, Iowa, USA from 2006 to 2010. The source of this data is Ames, Iowa Assessor's Office.

This data set comprises of 2930 data samples/rows of each containing 79 columns aka fields aka features aka dimensions. Additionally, some of the column values might be missing in the rows.

The data set is a classic mix of nominal/categorical values, discrete values, continuous values, and ordinal values for many input features/dimensions.

Provided files for this project:
-   Ames Housing Data – *ames.csv*
-   Project shell code file - *AAM-IPL-Wk-1-PCA-Ames-Housing-Shell-Code.ipynb*

The objective of the project is to:
-   Print the number of columns and number of rows/data samples/records in the provided data set
-   Identify numerical and categorical columns and print them
-   Fill the missing numeric values with mean of the respective column and categorical/ordinal columns with "Missing" category value
-   Scale numerical feature and encode categorical features and apply this transformation on the data set
-   Calculate cumulative explained variance up to the number of original features (79)
-   Plot explained variance ratio and cumulative explained variance ratio against principal components
-   Set the threshold of 0.9 (90%) for the cumulative explained variance to find optimal principal components (hyper parameter) and print the count of principal components
-   Plot heatmap of principal components correlation
-   Pair plot the first 5 principal components
-   Generate the PDF from the Jupyter file (code and output) and upload to the assignment

C-501, Salarpuria Serenity, 5th Main, Sector 7, HSR Layout, Bengaluru 560102 KA India
Mobile: +91 97012 22130, Email: vmelachervu@gmail.com, vmela23@iitk.ac.in, Website: www.linkedin.com/in/vmelachervu

Page 1 of 3

Venkateswar Reddy M.
CTO, Brillium Technologies
"...dare to dream; care to win..."

Date: **10-Sep-24**

Project Overview, Implementation and Submission Timeline

Sr. No.	Date	Project Topic	Comments
10-09-2024 – Tuesday 8:00 PM – PCA Project Details Announcement – Topic, Data Set, Shell Code etc. Announcement Channels – Google Class, Industry Projects WhatsApp Group.			
2	14-09-2024 - Saturday Duration: 1.5 Hrs	Principal Component Analysis (PCA) – Overview/Recap, Interactive Q&A	Online – Google Class
3	15-09-2024 - Sunday Duration: 1.5 Hrs	Principal Component Analysis (PCA) – Implementation, Output Demonstration, Interactive Q&A	Online – Google Class
15-09-2024 – Sunday 11:59 PM - Deadline to upload the project code submission by all students in Google Class.			

Guest Lecture Timings:
Saturdays: 10:30 AM IST – 12:00 Noon IST
Sundays: 10:30 AM IST – 12:00 Noon IST
Mondays: 6:30 PM IST – 8:00 PM IST

Development Environment
- Computing Language – Python
- IDE – Visual Studio Code with Jupyter Notebook

Instructor

Instructor	
Venkateswar Reddy Melachervu (alumnus of GPREC, ECE Class of '92) CTO, Brillium Technologies, Bengaluru Email: **venkat.reddy.gf@gprec.ac.in** Profile: LinkedIn	Visiting Faculty

Coordination
All the activities of this programme – lecture venues, weekly projects details announcements, general announcements, changes in lecture timings, etc. will be coordinated by CSM faculty member Sri V.Suresh.

Channels of Communication and Announcements
- Google Classroom
- Whatsapp group - **Applied AI & ML Industry Projects Lab**
- Emails (Strictly GPREC email addresses only)

Programme Coordinator	
Prof. V.Suresh Email: **vsuresh.ecs@gprec.ac.in**	Faculty Member, CSM

Reference Books
- Pattern Recognition and Machine Learning by Chris Bishop, 2006 – PDF Link

C-501, Salarpuria Serenity, 5th Main, Sector 7, HSR Layout, Bengaluru 560102 KA India
Mobile: +91 97012 22130, Email: vmelachervu@gmail.com, vmela23@iitk.ac.in, Website: www.linkedin.com/in/vmelachervu

Page 2 of 3

Venkateswar Reddy M.
CTO, Brillium Technologies
"...dare to dream; care to win..."

Date: **10-Sep-24**

- Machine Learning using Python by Manaranjan Pradhan and U Dinesh Kumar, Wiley 2019 – PDF Link

Policies

- **Attendance:** All sessions are expected to be attended by all the enrolled students. In case of inability to attend, prior information is expected to be provided by the student to the coordinator with a copy to the visiting faculty
- **Project Submissions:** Duly completed projects (Jupyter NB file and a PDF of the Jupyter NB file) are expected to be submitted through google class prior to the deadline. In case of inability to complete due various unforeseen circumstance, students are expected to seek extension for the submission deadline.
- **Academic Integrity:** Students are expected to uphold the highest standards of academic integrity in all assignments for the Applied AI & ML Industry Projects Lab. Each assignment must be the student's own work, and all sources and collaborators must be properly acknowledged. By submitting their completed project source code, students confirm that they have adhered to this integrity policy and completed their work in an honest and ethical manner.

Additional Information

For students interested in engaging with special projects in the field of Gen AI, please reach out to the visiting faculty at **venkat@brillium.in** for further details and opportunities.

Contact Information

For any questions or concerns or further details on this programme, please contact **Program Coordinator** during office hours or via email.

--- End of the Document---

C-501, Salarpuria Serenity, 5th Main, Sector 7, HSR Layout, Bengaluru 560102 KA India
Mobile: +91 97012 22130, Email: vmelachervu@gmail.com, vmela23@iitk.ac.in, Website: www.linkedin.com/in/vmelachervu

Page 3 of 3

"...dare to dream; care to win..."