Report of the Faculty Development Program on Data Analytics using KNIME Tool

Date: 11.02.2023

Venue : Mobile and IoT Lab, PSGiTech

The Department of Computer Science and Engineering in association with PSG Center for Academic Research and Excellence organized Faculty Development Program on Data Analytics using KNIME Tool scheduled on 11.02.2023, 9.30 am to 4.30 pm. **Dr G Chandramohan, Principal, PSGiTech** welcomed the resource person **Dr B Somasundaram, Head Digital Transformation**, **ELGi Equipment Ltd**. He is having 20 years of experience in analytics, digital technologies, and cutting-edge technologies in top firms like Siemens and ELGi equipment.

Dr R Manimegalai, Head of CSE, PSGiTech introduced the resource person and welcomed the gathering. She emphasized the importance of conducting a Faculty Development Program on Data Analytics using KNIME Tool.
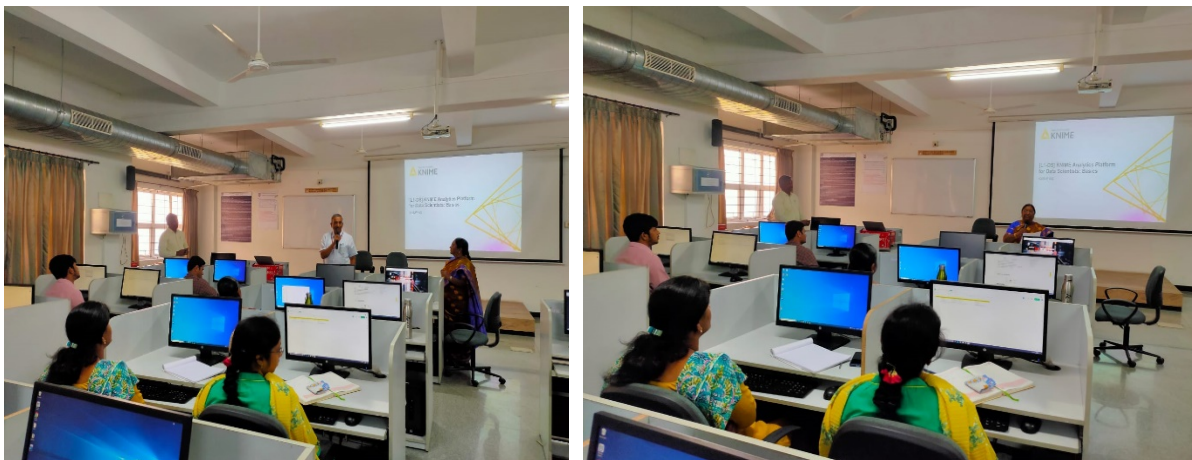
The FDP is coordinated by **Dr.B.Gomathi, ASP/CSE** and **Dr.S.Lokesh, ASP/CSE**. The FDP was well-attended by 27 faculty members from various colleges in Tamil Nadu.

The resource person **Dr B Somasundaram** enlightened the audience with **KNIME Analytics Package,** that offers a complete platform for end to end data science, from creating analytic models, to deploying them and sharing insights within the organization, through data apps and services KNIME Analytics Platform is open source software that allows users to access, blend, analyze, and visualize data, without any coding. It slow-code, no-code interface offers an easy introduction for beginners, and an advanced data science set of tools for experienced users.

KNIME is a low-code data science and data preparation platform that makes understanding data and designing analytic work flows accessible to everyone. The KNIME suite includes two tools: KNIME Analytics Platform is a desktop-based tool where analysts and developers construct workflows.
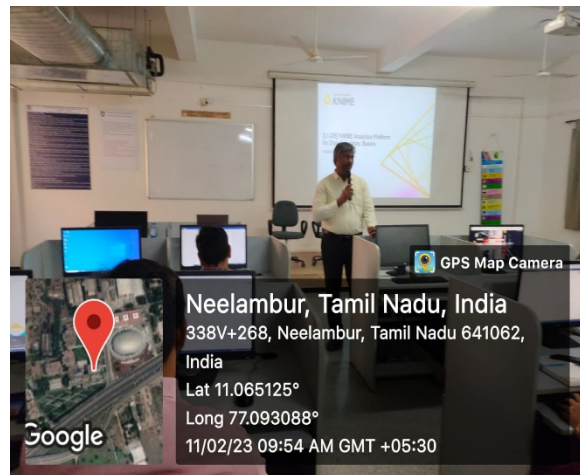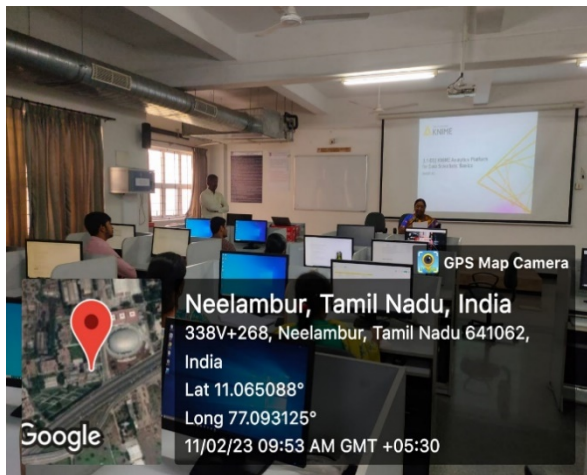
He elaborately discussed the following topics with hands on training:

1. **KNIME Analytics Platform**

   KNIME Analytics Platform is an open source software with an intuitive, visual interface that lets you build analyses of any complexity level – from automating spreadsheets to ETL to machine learning. Users can access, blend, analyze, and visualize data, without any coding, or integrate their favorite tools and libraries as needed.

2. **KNIME Community Hub**

    With KNIME Community Hub, users can browse and learn from thousands of working examples of data science solutions, benefiting from community contributions to upskill and dive deeper into the discipline of data science. Small groups or teams of users can also share and collaborate on solutions in their dedicated, private spaces.

3. **Shape your Data**

    Derive statistics, including mean, quantiles, and standard deviation, or apply statistical tests to validate a hypothesis. Integrate dimensions reduction, correlation analysis, and more into your workflows. Aggregate, sort, filter, and join data either on your local machine, in-database, or in distributed big data environments. Clean data through normalization, data type conversion, and missing value handling. Detect out of range values with outlier and anomaly detection algorithms. Extract and select features (or construct new ones) to prepare your dataset for machine learning with genetic algorithms, random search or backward- and forward feature elimination. Manipulate text, apply formulas on numerical data, and apply rules to filter out or mark samples.

4. **Leverage Machine Learning & AI**

    Build machine learning models for classification, regression, dimension reduction, or clustering, using advanced algorithms including deep learning, tree-based methods, and logistic regression. Access any popular ML library like TensorFlow, Keras, H2O, and more and cutting-edge techniques in text mining, image processing, and more. Validate models by applying performance metrics including Accuracy, R2, AUC, and ROC. Perform cross validation to guarantee model stability. Explain machine learning models with LIME, Shap/Shapley values. Understand model predictions with the interactive partial dependence/ICE plot.

5. **Discover and Share Data Insights**

    Visualize data with classic (bar chart, scatter plot) as well as advanced charts (parallel coordinates, sunburst, network graph, heat map) and customize them to your needs. Export reports as PDF, PowerPoint, or other formats for presenting results to stakeholders. Store processed data or analytics results in many common file formats or databases.

Dr.B.Gomathi, co-ordinator of the FDP proposed the formal vote of thanks. He extended his heartfelt thanks to resource person and faculty members who participated and helped in making this FDP great success. Later, group photo were taken along with participants.