**Project preparations**

1. You will have to prepare a project report with the following content:
   1. Table of content
   2. Table of figures/graphs
   3. Abstract
   4. Chapters (Introduction, Software-libraries used, Algorithm etc.)
   5. Conclusion
   6. Complete code
   7. Font style: Times New Roman
   8. Font size: Chapter heading: 16 (bold), Topics: 14 (bold), content: 14
   9. Margin: 1” on all sides
   10. The front page will contain the logo of your college and the company - Knowledge solutions India.
2. You will have to prepare a presentation of each individual candidate, explaining their parts in the project and will have to present it.
3. The report will be submitted to us and your respective college(if they ask) . In the companies copy the reports will be submitted in my name:

Gurvansh Singh

M.Tech

Knowledge Solutions India

1. Time period for the project is 14-15 days, the code file should also be submitted.

**Project – 3**

**NLP – Review classification**

Content

The dataset contains reviews and their classification as good and bad reviews.  
The parameters included are:

1. Reviews
2. Label (0 – bad review and 1 – good review)

Create a ML models to predict the Review and plot the confusion matrix(hint: sklearn.metrics -> plot\_confusin\_matrix <https://scikit-learn.org/stable/modules/generated/sklearn.metrics.plot_confusion_matrix.html>) for all the classifier models and also calculate the accuracy score for each. Build the following models:

1. Support Vector Machines (SVM)
2. K – nearest neighbor classifier (K-NN)
3. SVM with PCA

Perform a very detailed NLP on the dataset and try to increase the accuracy of the models to atleast 90%. And plot the classification output.