Project Synopsis

on

# YOUCAN- PLANNER AND TIME MANAGER

Submitted as a part of course curriculum for

**Bachelor of Technology**

in

## Computer Science



**Submitted by**

**SNEHA SHARMA(1900290120116)**

**TEJASVI CHAUDHARY(1900290120120)**

**Under the Supervision of**

Prof. Pradeep Tyagi

**KIET Group of Institutions, Ghaziabad Department of Computer Science Dr. A. P. J. Abdul Kalam Technical University**

**2021-2022**

## DECLARATION

We hereby declare that this submission is our work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgement has been made in the text.

Signature of Students Name:

Roll No.:

Date:

## CERTIFICATE

This is to certify that Project Report entitled “**Title**” which is submitted by **Student Name** in partial fulfilment of the requirement for the award of degree B. Tech. in Department of Computer Science of Dr A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

**Date: Supervisor Signature**

Supervisor Name (Designation)

## ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the synopsis of the B.Tech Mini Project undertaken during B.Tech. Third Year. We owe a special debt of gratitude to GUIDE NAME WITH DESIGNATION, Department of Computer Science, KIET Group of

Institutions, Delhi- NCR, Ghaziabad, for his/her constant support and guidance throughout the course of our work. His/Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his/her cognizant efforts that our endeavours have seen the light of the day.

We also take the opportunity to acknowledge the contribution of Dr. P. K Singh, Head of the Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

Last but not the least, we acknowledge our friends for their contribution to the completion of the project.

Signature:

Date :

Name :

Roll No:

## ABSTRACT

◦ With the help of this project, our team wants to provide society with a time saving as well as time managing website.

◦ . Do you find that there is never enough time for everything? If so, it is probably because you are not scheduling your time effectively. This happens when you just run constantly, going from one thing to the next, without really thinking about everything as a whole. Using a planner allows you to schedule each event, appointment, errand, and task, so that you know what to expect and don’t run out of time. Set specific due dates and deadlines for everything to help you stay on track, and have extra time left for yourself and your family.

◦ Planner is a web application built in MERN stack technologies. It constitutes 4 services, a ◦ calendar, events, notes and task tracker.

◦ This whole project is built on a node runtime environment.

◦ To build the frontend, the most popular frontend framework has been used - React JS

◦ with a number of libraries prepared for react.

◦ For CSS, the SASS preprocessor has been used.

◦ For managing the backend efficiently, the most efficient framework :Express JS has been

◦ used.

◦ All the data will be saved in a database managed by MongoDB.

|  |  |
| --- | --- |
| **TABLE OF CONTENTS** | Page No. |
| TITLE PAGE .................................................................................................................... | i |
| DECLARATION .............................................................................................................. | ii |
| CERTIFICATE …........................................................................................................... | iii |
| ACKNOWLEDGEMENT.................................................................................................. | iv |
| ABSTRACT...................................................................................................................... | v |
| LIST OF FIGURES ......................................................................................................... | vi |

|  |  |
| --- | --- |
| LIST OF ABBREVIATIONS ………..…………………………………………………. vii | |
| CHAPTER 1 INTRODUCTION | 1-n |
| 1.1. Introduction ……………………................................................... | 1 |
| 1.2 Problem Statement .……………………....................................... |  |
| 1.2. Objective………………………………………………………… | 2 |
| 1.3. Scope…………………………………………………………….. | 3 |
| CHAPTER 2 LITERATURE REVIEW…………………………………………….... | 7-p |
| CHAPTER 3 PROPOSED METHODOLOGY …………………………………........ | 8-m |
| 3.1 Flowchart |  |
| 3.2 Algorithm Proposed | 10 |
| CHAPTER 4 TECHNOLOGY USED ………..………………………..………………. | 12 |
| CHAPTER 5 DIAGRAMS …...........................................................................................  CHAPTER 6 CONCLUSION …....................................................................................... REFERENCES….............................................................................................................. | |

**The content of the synopsis must be divided into following headings depending upon your project**

1. INTRODUCTION

With the help of this project, our team wants to provide society with a time saving as well as time managing website.

1. PROBLEM STATEMENT

Do you find that there is never enough time for everything? If so, it is probably because you are not scheduling your time effectively. This happens when you just run constantly, going from one thing to the next, without really thinking about everything as a whole. Using a planner allows you to schedule each event, appointment, errand, and task, so that you know what to expect and don’t run out of time. Set specific due dates and deadlines for everything to help you stay on track, and have extra time left for yourself and your family.

1. OBJECTIVES

Manage , capture and edit your daily events , notes and tasks.

1. LITERATURE REVIEW

◦ 1. TITLE: MLife: a lite framework for machine learning lifecycle initialization.

◦ SUMMARY: Machine learning (ML) lifecycle is a cyclic process to build an efficient ML system. Though a lot of commercial and community (non-commercial) frameworks have been proposed to streamline the major stages in the ML lifecycle, they are normally overqualified and insufficient for an ML system in its nascent phase. Driven by real-world experience in building and maintaining ML systems, we find that it is more efficient to initialize the major stages of ML lifecycle first for trial and error, followed by the extension of specific stages to acclimatize towards more complex scenarios. For this, we introduce a simple yet flexible framework, MLife, for fast ML lifecycle initialization. This is built on the fact that data flow in MLife is in a closed loop driven by bad cases, especially those which impact ML model performance the most but also provide the most value for further ML model development—a key factor towards enabling enterprises to fast track their ML capabilities. Better yet, MLife is also flexible enough to be easily extensible to more complex scenarios for future maintenance. For this, we introduce two real-world use cases to demonstrate that MLife is particularly suitable for ML systems in their early phases.

◦ 2. TITLE: Sparse classifcation: a scalable discrete optimization perspective

◦ SUMMARY: We formulate the sparse classifcation problem of n samples with p features as a binary convex optimization problem and propose a outer-approximation algorithm to solve it exactly. For sparse logistic regression and sparse SVM, our algorithm fnds optimal solutions for n and p in the 10,000 s within minutes. On synthetic data our algorithm achieves perfect support recovery in the large sample regime. Namely, there exists an n0 such that the algorithm takes a long time to fnd an optimal solution and does not recover the correct support for n < n0, while for n ⩾ n0, the algorithm quickly detects all the true features, and does not return any false features. In contrast, while Lasso accurately detects all the true features, it persistently returns incorrect features, even as the number of observations increases. Consequently, on numerous real-world experiments, our outer-approximation algorithms returns sparser classifers while achieving similar predictive accuracy as Lasso. To support our observations, we analyze conditions on the sample size needed to ensure full support recovery in classifcation.

|  |  |
| --- | --- |
| ◦ | 3. TITLE: RB-CCR: Radial-Based Combined Cleaning and Resampling algorithm for imbalanced data classification. |
| ◦ | SUMMARY: Real-world classifcation domains, such as medicine, health and safety, and fnance, often exhibit imbalanced class priors and have asynchronous misclassifcation costs. In such cases, the classifcation model must achieve a high recall without signifcantly impacting precision. Resampling the training data is the standard approach to improving classifcation performance on imbalanced binary data. However, the state-of-the-art methods ignore the local joint distribution of the data or correct it as a post-processing step. This can causes sub-optimal shifts in the training distribution, particularly when the target data distribution is complex. In this paper, we propose  Radial-Based Combined Cleaning and Resampling (RB-CCR). RB-CCR utilizes the concept of class potential to refne the energy-based resampling approach of CCR. In particular, RB-CCR exploits the class potential to accurately locate sub-regions of the data-space for synthetic oversampling. The category sub-region for oversampling can be specifed as an input parameter to meet domain-specifc needs or be automatically selected via cross-validation. Our 5 × 2 cross-validated results on 57 benchmark binary datasets with 9 classifers show that RB-CCR achieves a better precision-recall trade-of than CCR and generally out-performs the state-of-the-art resampling methods in terms of AUC and G-mean. |
| ◦ | 4. TITLE: Deep learning and multivariate time series for cheat detection in video games. |
| ◦ | SUMMARY: Online video games drive a multi-billion dollar industry dedicated to maintaining a competitive and enjoyable experience for players. Traditional cheat detection systems struggle when facing new exploits or sophisticated fraudsters. More advanced solutions based on machine learning are more adaptive but rely heavily on in-game data, which means that each game has to develop its own cheat detection system. In this work, we propose a novel approach to cheat detection that doesn’t require in-game data. Firstly, we treat the multimodal interactions between the player and the platform as multivariate time series. We then use convolutional neural networks to classify these time series as corresponding to legitimate or fraudulent gameplay. Our models achieve an average accuracy of respectively 99.2% and 98.9% in triggerbot and aimbot (two widespread cheats), in an experiment to validate the system’s ability to detect cheating in players never seen before. Because this approach is based solely on player behavior, it can be applied to any game or input method, and even various tasks related to modeling human activity |

|  |  |
| --- | --- |
| ◦ | 5. TITLE: A unified framework for closed-form nonparametric regression, classification, preference and mixed problems with Skew Gaussian Processes. |
| ◦ | SUMMARY: Skew-Gaussian Processes (SkewGPs) extend the multivariate Unifed Skew-Normal distributions over fnite dimensional vectors to distribution over functions. SkewGPs are more general and fexible than Gaussian processes, as SkewGPs may also represent asymmetric distributions. In a recent contribution, we showed that SkewGP and probit likelihood are conjugate, which allows us to compute the exact posterior for non-parametric binary classifcation and preference learning. In this paper, we generalize previous results and we prove that SkewGP is conjugate with both the normal and afne probit likelihood, and more in general, with their product. This allows us to (i) handle classifcation, preference, numeric and ordinal regression, and mixed problems in a unifed framework; (ii) derive closed-form expression for the corresponding posterior distributions. We show empirically that the proposed framework based on SkewGP provides better performance than Gaussian processes in  active learning and Bayesian (constrained) optimization. These two tasks are fundamental for design of experiments and in Data Science |
| ◦ | 6. TITLE: Loss aware post-training quantization. |
| ◦ | SUMMARY: Neural network quantization enables the deployment of large models on resource-constrained devices. Current post-training quantization methods fall short in terms of accuracy for INT4 (or lower) but provide reasonable accuracy for INT8 (or above). In this work, we study the efect of quantization on the structure of the loss landscape. We show that the structure is fat and separable for mild quantization, enabling straightforward post-training quantization methods to achieve good results. We show that with more aggressive quantization, the loss landscape becomes highly non-separable with steep curvature, making the selection of quantization parameters more challenging. Armed with this understanding, we design a method that quantizes the layer parameters jointly, enabling signifcant accuracy improvement over current post-training quantization methods. |

|  |  |
| --- | --- |
| ◦ | 7. TITLE: Misalignment problem in matrix decomposition with missing values. |
| ◦ | SUMMARY: Data collection within a real-world environment may be compromised by several factors such as data-logger malfunctions and communication errors, during which no data is collected. As a consequence, appropriate tools are required to handle the missing values when analysing and processing such data. This problem is often tackled via matrix decomposition. While it has been successfully applied in a wide range of applications, in this work we report an issue that has been neglected in literature and “degenerates” the quality of the imputations obtained by matrix decomposition in multivariate time-series (with smooth evolution). Briefy, the problem consists of the misalignment of the matrix decomposition result: the missing values imputations fall within an incorrect range of values and the transitions between observed and imputed values are not smooth. We address this problem by proposing a post-processing alignment strategy. According to our experiments, the post-processing adjustment substantially improves the accuracy of the imputations (when the misalignment occurs). Moreover, the results also suggest that the misalignment occurs mostly when dealing with a small number of time-series due to lack of generalization ability |
| ◦ | 8. TITLE: Tensor decision trees for continual learning from drifting data streams. |
| ◦ | SUMMARY: Data stream classifcation is one of the most vital areas of contemporary machine learning, as many real-life problems generate data continuously and in large volumes. However, most of research in this area focuses on vector-based representations, which are unsuitable for capturing properties of more complex multidimensional structures, such as images and video sequences. In this paper, we propose a novel methodology for learning adaptive decision trees from data streams of tensors. We introduce Chordal Kernel Decision Tree for continual learning from tensor data streams. In order to maintain the tensor characteristics, we propose to train and update classifers in the kernel space designed to work with tensor representation. We use chordal distance to compute similarities between tensors and then apply it as a new feature space in which decision trees are trained. This allows for a direct decision tree induction on tensors. In order to accommodate the streaming and drifting nature of data, we propose a concept drift detection scheme based on tensor representation. It allows us to reconstruct the kernel feature space every time when change is detected. The proposed approach allows for fast and efcient induction of decision trees on streaming data with tensor representation. Experimental study, conducted on 4 realworld and 52 artifcial large-scale tensor data streams, shows that using the native tensor feature space leads to more accurate classifcation than outperforms the vectorized representations. |

|  |  |
| --- | --- |
| ◦ | 9. TITLE: Symbolic DNN-Tuner. |
| ◦ | SUMMARY: Hyper-Parameter Optimization (HPO) occupies a fundamental role in Deep Learning systems due to the number of hyper-parameters (HPs) to be set. The state-of-the-art of HPO methods are Grid Search, Random Search and Bayesian Optimization. The frst two methods try all possible combinations and random combination of the HPs values, respectively. This is performed in a blind manner, without any information for choosing the new set of HPs values. Bayesian Optimization (BO), instead, keeps track of past results and uses them to build a probabilistic model mapping HPs into a probability density of the objective function. Bayesian Optimization builds a surrogate probabilistic model of the objective function, fnds the HPs values that perform best on the surrogate model and updates it with new results. In this paper, we improve BO applied to Deep Neural Network (DNN) by adding an analysis of the results of the network on training and validation sets. This analysis is performed by exploiting rule-based programming, and in particular by using Probabilistic Logic Programming. The resulting system, called Symbolic DNN-Tuner, logically evaluates the results obtained from the training and the validation phase and, by applying symbolic tuning rules, fxes the network architecture, and its HPs, therefore improving performance. We also show the efectiveness of the proposed approach, by an experimental evaluation on literature and real-life datasets. |
| ◦ | 10. TITLE: Reinforcement learning for robotic manipulation using simulated locomotion demonstrations. |
| ◦ | SUMMARY: Mastering robotic manipulation skills through reinforcement learning (RL) typically requires the design of shaped reward functions. Recent developments in this area have demonstrated that using sparse rewards, i.e. rewarding the agent only when the task has been successfully completed, can lead to better policies. However, state-action space exploration is more difcult in this case. Recent RL approaches to learning with sparse rewards have leveraged high-quality human demonstrations for the task, but these can be costly, time consuming or even impossible to obtain. In this paper, we propose a novel and efective approach that does not require human demonstrations. We observe that every robotic manipulation task could be seen as involving a locomotion task from the perspective of the object being manipulated, i.e. the object could learn how to reach a target state on its own. In order to exploit this idea, we introduce a framework whereby an object locomotion policy is initially obtained using a realistic physics simulator. This policy is then used to generate auxiliary rewards, called simulated locomotion demonstration rewards (SLDRs), which enable us to learn the robot manipulation policy. The proposed approach has been evaluated on 13 tasks of increasing complexity, and can achieve higher success rate and faster learning rates compared to alternative algorithms. SLDRs are especially benefcial for tasks like multi-object stacking and non-rigid object manipulation. |

|  |  |
| --- | --- |
| ◦ | 11. TITLE: Stronger data poisoning attacks break data sanitization defenses. |
| ◦ | SUMMARY: Machine learning models trained on data from the outside world can be corrupted by data poisoning attacks that inject malicious points into the models’ training sets. A common defense against these attacks is data sanitization: frst flter out anomalous training points before training the model. In this paper, we develop three attacks that can bypass a broad range of common data sanitization defenses, including anomaly detectors based on nearest neighbors, training loss, and singular-value decomposition. By adding just 3% poisoned data, our attacks successfully increase test error on the Enron spam detection dataset from 3 to 24% and on the IMDB sentiment classifcation dataset from 12 to 29%. In contrast, existing attacks which do not explicitly account for these data sanitization defenses are defeated by them. Our attacks are based on two ideas: (i) we coordinate our attacks to place poisoned points near one another, and (ii) we formulate each attack as a constrained optimization problem, with constraints designed to ensure that the poisoned points evade detection. As this optimization involves solving an expensive bilevel problem, our three attacks correspond to diferent ways of approximating this problem, based on infuence functions; minimax duality; and the Karush–Kuhn–Tucker (KKT) conditions. Our results underscore the need to develop more robust defenses against data poisoning attacks. |
| ◦ | 12. TITLE: Learning any memory-less discrete semantics for dynamical systems represented by logic programs. |
| ◦ | SUMMARY: Learning from interpretation transition (LFIT) automatically constructs a model of the dynamics of a system from the observation of its state transitions. So far the systems that LFIT handled were mainly restricted to synchronous deterministic dynamics. However, other dynamics exist in the feld of logical modeling, in particular the asynchronous semantics which is widely used to model biological systems. In this paper, we propose a modeling of discrete memory-less multi-valued dynamic systems as logic programs in which a rule represents what can occur rather than what will occur. This modeling allows us to represent non-determinism and to propose an extension of LFIT to learn regardless of the update schemes, allowing to capture a large range of semantics. We also propose a second algorithm which is able to learn a whole system dynamics, including its semantics, in the form of a single propositional logic program with constraints. We show through theoretical results the correctness of our approaches. Practical evaluation is performed on benchmarks from biological literature |

|  |  |
| --- | --- |
| ◦ | 13. TITLE: Bimodal variational autoencoder for audiovisual speech recognition. |
| ◦ | SUMMARY: Multimodal fusion is the idea of combining information in a joint representation of multiple modalities. The goal of multimodal fusion is to improve the accuracy of results from classifcation or regression tasks. This paper proposes a Bimodal Variational Autoencoder (BiVAE) model for audiovisual features fusion. Reliance on audiovisual signals in a speech recognition task increases the recognition accuracy, especially when an audio signal is corrupted. The BiVAE model is trained and validated on the CUAVE dataset. Three classifers have evaluated the fused audiovisual features: Long-short Term Memory, Deep Neural Network, and Support Vector Machine. The experiment involves the evaluation of the fused features in the case of whether two modalities are available or there is only one modality available (i.e., cross-modality). The experimental results display the superiority of the proposed model (BiVAE) of audiovisual features fusion over the state-of-the-art models by an average accuracy diference  3.28% and 13.28% for clean and noisy, respectively. Additionally, BiVAE outperforms the state-of-the-art models in the case of crossmodality by an accuracy diference  2.79% when the only audio signal is available and 1.88% when the only video signal is available. Furthermore, SVM satisfes the best recognition accuracy compared with other classifers. |
| ◦ | 14. TITLE: Embedding and extraction of knowledge in tree ensemble classifiers. |
| ◦ | SUMMARY: The embedding and extraction of knowledge is a recent trend in machine learning applications, e.g., to supplement training datasets that are small. Whilst, as the increasing use of machine learning models in security-critical applications, the embedding and extraction of malicious knowledge are equivalent to the notorious backdoor attack and defence, respectively. This paper studies the embedding and extraction of knowledge in tree ensemble classifers, and focuses on knowledge expressible with a generic form of Boolean formulas, e.g., point-wise robustness and backdoor attacks. For the embedding, it is required to be preservative (the original performance of the classifer is preserved), verifable (the knowledge can be attested), and stealthy (the embedding cannot be easily detected). To facilitate this, we propose two novel, and efective embedding algorithms, one of which is for blackbox settings and the other for white-box settings. The embedding can be done in PTIME. Beyond the embedding, we develop an algorithm to extract the embedded knowledge, by reducing the problem to be solvable with an SMT (satisfability modulo theories) solver. While this novel algorithm can successfully extract knowledge, the reduction leads to an NP computation. Therefore, if applying embedding as backdoor attacks and extraction as defence, our results suggest a complexity gap (P vs. NP) between the attack and defence when working with tree ensemble classifers. We apply our algorithms to a diverse set of datasets to validate our conclusion extensively. |

|  |
| --- |
| ◦ 15. TITLE: A stochastic approach to handle resource constraints as knapsack problems in ensemble pruning.  ◦ SUMMARY: Ensemble-based methods are highly popular approaches that increase the accuracy of a decision by aggregating the opinions of individual voters. The common point is to maximize accuracy; however, a natural limitation occurs if incremental costs are also assigned to the individual voters. Consequently, we investigate creating ensembles under an additional constraint on the total cost of the members. This task can be formulated as a knapsack problem, where the energy is the ensemble accuracy formed by some aggregation rules. However, the generally applied aggregation rules lead to a nonseparable energy function, which takes the common solution tools—such as dynamic programming—out of action. We introduce a novel stochastic approach that considers the energy as the joint probability function of the member accuracies. This type of knowledge can be efciently incorporated in a stochastic search process as a stopping rule, since we have the information on the expected accuracy or, alternatively, the probability of fnding more accurate ensembles. Experimental analyses of the created ensembles of pattern classifers and object detectors confrm the efciency of our approach over other pruning ones. Moreover, we propose a novel stochastic search method that better fts the energy, which can be incorporated in other stochastic strategies as well.         1. CONCLUSION WITH RESULT   Manage , capture and edit your daily events , notes and tasks.     1. References <https://fonts.google.com/><https://photos.google.com/><https://keep.google.com/><https://material.io/> |

|  |
| --- |
|  |