NIKHIL AGGARWAL

Senior Undergraduate
Department of Computer Science and Engineering
Indian Institute of Technology, Kanpur, India
E-mail: nikhil[AT]iitk.ac.in , nikhil23393[AT]gmail.com

A-305, Hall 1
IIT Kanpur-208016
Phone: + (91) 7607481193

EDUCATION

Year	Degree	Institution	Performance
2014	B.Tech, Computer Science and Engineering	IIT Kanpur	**8.5/10.0***
2010	Class 12 : CBSE Board	Apeejay School, Sheikh Sarai, New Delhi	92.2%
2008	Class 10 : ICSE Board	Christu Jyoti Convent School,Baghpat	93.3%

*after completion of 6 semesters

SCHOLASTIC ACHIEVEMENTS

AWARDS	• Awarded **Academic Excellence Award** for academic year 2010-11, awarded to top few students in the department • Awarded Certificate of merit for being placed in top 1% in **National Physics Olympiad** 2010 held by Indian Association of Physics Teachers and qualified for next level exam • Awarded Certificate of merit for being placed in top 1% in **National Chemistry Olympiad** 2010 held by Indian Association of Chemistry Teachers and qualified for next level exam • Awarded **Best Project Award** in Introduction to Manufacturing Processes course among 60 other projects • Awarded **Second Best Project Award** for making paddle-boat in TA201 course among 65 other projects
SCHOLARSHIPS	• Awarded **Kishore Vaigyanik Protsahan Yojana (KVPY)** scholarship for the year 2009-10 (***top 90 students*** across India) • Awarded **CBSE Merit Scholarship** under AIEEE 2010. Awarded to 332 students from a pool of 11Lac students • Awarded **Merit-cum-Means Scholarship** for meritorious performance for three consecutive years since 2010 • Awarded Certificate of excellence in **IGSC Scholarship Examination** held at National level
RANKS	• Achieved an **All India Rank(AIR) 229** in **IIT-JEE'10** in which nearly 5Lac students appeared. **Percentile-99.95%** • Secured **AIR 262 Rank** in AIEEE'10 in which nearly 1,100,000 students appeared. **Percentile- 99.98%** • Secured **AIR 6** in UPTU Examination 2010. **Percentile- 99.99%**

INTERNSHIPS / WORK EXPERIENCE

Content Monitoring for Wal-Mart affiliates (WM Global Technology Services India Pvt. Ltd, Bengaluru) *(May-July 2013)*

Technology used: coding in JAVA, eclipse IDE, SBT project, Machine Learning

- Made a binary classifier for Webpages: Given a URL Classify Webpage as Bad if it contains adult, racist, communal content otherwise Good. Also Display Wal-Mart adds only on good pages.
- Generated probability of occurrence of a word in bad page and good page. Based on these probabilities a score (word probability score) was calculated for the page. Also we were using an API that returned the concepts for the page. Generated a concept vs. concepts Matrix and a score (matrix score) was assigned to the webpage based on this matrix.
- Used logistic regression to classify webpage as Bad or Good using word probability score and matrix score. Overall precision increased from 70% to 95%. Also recall for Bad pages increased from 57% to 95%.
- We also did classification just based on URL and not the content of the Webpage by looking at the structure of the URL.

Walmart Display its advertisement on various sites. But it doesn't want to show its ad on bad pages which contain Adult, Racist, Communal, etc. content. For this it needs to monitor the pages on which it could Display its ad. Overall project was to build a content monitoring module which could classify pages into two categories:

1. one which adhere to walmart policies and we could display our ad on those pages
2. other which Doesn't adhere to walmart policies and we shouldn't Display our ads on it

We were already using a classifier but were not sure about the performance. First part was to build a new project and implement the current classifier and then do a performance analysis. Current system had two parts:

1. Adult Word Percentage analysis: It calculates the percentage of adult words in the web page content and if it is greater than some threshold then classify page as Bad page otherwise good.
2. Alchemy Concept Analysis: Alchemy API is a Natural Language Processing service which provides different sematic analysis of data. We used concept tagging in which alchemy attached some concepts to the page and confidence (0-1) of that concept for that page.

We would like to have a system which has:

1. High Recall for Bad pages
2. High precision

Current System had poor performance with recall for bad page as 48.54% and precision as 75.50%. I built a Data Set and analysed different Threshold for both Adult Word Percentage analysis and Alchemy Concept Analysis and increased the performance to recall for bad pages 76.70 % and precision 94.24%. I further tried to improve the performance by using these features:

1. Word probability score: I generated the probability of occurrence of word in Bad page (Pb) and good page (Pg) by training on the content of 20,000 Bad pages and 20,000 Good Pages. This training set was built from 3 lakh Bad URL and 2.5 lakh Good URLs after filtering URLs based on some criteria Then a score was assigned to each pages based on these probability.
2. Concept vs. concept Matrix score: Generated Concept vs. concept Matrix to capture the context in which concept is occurring. Again Matrix was generated on the above mentioned Dataset.
3. Some other features like Bad concept confidence sum, Bad concept count, etc. were also introduced.

I generated these features score for the data set and ran logistic regression and OneR rule to Train model on some training set and then put the parameter generated in the actual code. I tried different combinations of feature to remove noise and get the best results. Performance increased to

1. Recall for bad pages – 93%
2. Precision – 96%

I also implemented stemming of words using porter stemmer. But final code doesn't include it. Also I looked into classification based only on URLs. This doesn't include domain level classification but we looked at the structure of URL and tried to classify. Performance was quite good (recall around 70%). But all URLs caught on URL based classification were also caught by our new system. So that was removed from final Code.

Parallel Computing for Autonomous Vehicle Simulation (Carnegie Mellon University, USA) *(May-July 2012)*

Guide: Prof. Raghunathan Raj RajKumar, Department of Electrical and Computer Engineering, CMU

Technology used: coding in CUDA, GeForce GT530 Nvidia Graphic Card

- **Selected among 17 students** from across the globe for **ECE internship program** at Carnegie Mellon University
- Aim of the project was to make AutoSim work on GPU architecture. AutoSim is modular software that simulates Autonomous cars in hybrid environment consisting of real cars and autonomous cars. As AutoSim was coded in C++ it could only be used for simulating smaller number of cars and hence was not useful for inheriting real world map of cities.
- We implemented models on GPU and use parallelism of GPU while calling same models for different cars, hence enabling us to run simulation for thousands of car.
- Results were shown as execution time comparison between AutoSim running on single core CPU and GPU for different number of cars simulated. Also number of cars that could be simulated increased almost 50 times.

The goal of the project is to making AutoSim Work on GPU (device). AutoSim is software that simulates Autonomous car in hybrid environment i.e. real car and autonomous cars. AutoSim is been implemented in terms of models. Some models are:

1. Control Model
2. Position Model
3. State Model
4. Mobility Model
5. Path Tracker
6. Communication Model, etc.

Each model is actually a class and inheritable. So each car call these models to change its position, state, speed, direction, avoiding collision with other cars or road and many more things. AutoSim was coded in C++. Due to sequential execution of AutoSim it could only be used for simulating smaller number of cars and hence was not useful for inheriting real world map of cities in AutoSim and run it for few thousand cars.

We tried to implement models on GPU and use parallelism of GPU while calling same models for different cars. We implemented four models on GPU:

1. Position Model
2. Control Model
3. State Model
4. Path Tracker

As all models are not yet implemented on GPU we had to copy data from host (CPU) to device and again back to host after model is updated on GPU. This memory copy is bottle neck in execution. So for less number of cars we got poorer performance for GPU. But for larger number of car GPU was way better. We also made path tracker more efficient. Initially each car stored its path separately. But we have few static paths only. So instead of copying data for each car, we can just pass Path ID to car for accessing the path.

The future scope is to implement AutoSim completely on GPU and to add some other feature like dynamically generating path and including world model so that AutoSim can be used to generate real world simulation.

We collected data about time of execution for various models that were running on CPU and now are implemented on GPU for different number of cars and plotted them. Using GPU would be very useful not only for AutoSim but also for Autonomous Vehicles.
We used Qt-Creator to build project. We coded in CUDA and C++.

KEY ACADEMIC PROJECTS

Build Hindi and Bangla OCR and System for Crowd Sourcing: TO BE UPDATED *(August 2013 - ongoing)*

- Implement a crowd sourcing system that would take an image and perform OCR on it. User could also edit the output text
- Analyze the existing OCR systems for Hindi and Bangla Languages for their efficiency and robustness
- Implement a new OCR systems with new algorithm and do performance analysis based on correctness of word and character

Activity Manager: TO BE UPDATED *(August 2013 - ongoing)*

- The project aim to build a desktop application that runs in background and would keep track of all your activities on the laptop/PC and log all the data
- User could view log of any particular day and could analyze how much time he spent on computer and doing what activities. This would help user to increase his efficiency
- The log could also be synced with Google calendar so that if user is not on his laptop and have put something on his calendar then that activity would be logged for that particular interval
- User could also edit logs for his personal use and could also pause the application

Bill-Monk- Database Systems *(March-April 2013)*

Technology used: coding in PHP and used MySQL Database

- The project aimed to build a website where user could keep track of things he has borrowed from others and lend to others
- User could add shared bills like rent of house, itemized bill like shopping bill and could also do payments of his debts. Website will automatically distribute the amount. They could also shuffle the debt between friends
- User could make his own library of item classified as book, movie, music, clothes, others. He could borrow things from friend's library and lend things to his/her friends. He could also set a reminder to collect an item

Compiler for C# *(March-April 2013)*

Technology used: Lex, YACC and coding in C

- Designed a C# compiler which can generate code for SPIM architecture, with the following features:
  - Data types: Basic Data Types, composite data types, Identifiers, keywords, constants and literals
  - Operators: Arithmetic, logical, relational, bitwise and assignment operators
  - Statements: Assignment, conditional, switch, loop and comments
  - Functions: supporting parameters, return type and recursion
  - Object Oriented: Classes, Objects, Constructor, destructor, Access Modifiers (Public, private, internal and protected), Inheritance(Multiple and Multilevel)
- Compiler was completed in five stages: lexical analysis, syntax analysis, semantic analysis, intermediate code generation and final code generation. Intermediate code generated was three address code

Chess Playing Robot: Artificial Intelligence *(March-April 2013)*

Technology used: OpenCV, GNU chess, Lego mindstorm Robotics kit

- Designed and Build a robot using Lego kit that could play one on one chess with user
- Overhead Camera was used to detect the moves of the user and help robot to make its move
- GNU chess algorithm was used to calculate move for the robot. We could also set the difficulty level

Dynamic Graph Connectivity in poly-logarithmic worst case time*: Graph Algorithms* *(March-April 2013)*

Guide: Prof. Surendra Baswana, Department of computer Science and Engineering, IIT Kanpur

- Implemented the algorithm for dynamic graph connectivity in poly-logarithmic worst case time proposed by King, Mountjoy and Kapron. If algorithm say that there is a path from a to b then answer is correct but if it says the there is no such path then there could still be path from a to b.
- Analyzed the error in the answer for different number of vertices and edges in graph. Also compared the running time of this algorithm with trivial breath first search algorithm for different mixture of *insert(), delete() and query()* operations

Operating Systems *(Aug-Nov 2012)*

Technology used: coding in C and worked on *PINTOS* OS framework and *BOCHS* virtual machine

- The project aimed at providing various functionalities to **PINTOS**, instructional software that runs as secondary OS on Linux
- Implemented **POSIX message queue** with system calls like *open(), close(), unlink(), send()* and *receive()*
- Implemented an **indexed file-system** with direct, indirect and doubly indirect blocks so that files can grow and address the problem of external fragmentation. Also allowed for **hierarchical subdirectories** via system calls like *mkdir(), readdir(), chdir()* and provided for a **buffer cache**.
- Implemented virtual memory management via pure demand paging and shared memory via *open()* and *close().*
- Implemented **POSIX threads** and **scheduling algorithms** like First-Come-First-Serve, Round Robin and Priority based. Implemented system calls like ***fork()* and *exec().***

P2P File Sharing System *(Aug-Nov 2012)*

Technology used: coding in socket programming

- Implemented a Napster like P2P file sharing system where each user can **act as both server and client.**
- The system had a **central server** which keeps a log of the files shared by all the users who have currently logged in.
- Any user can request for a file to the central server which would return the name of the user who has shared that file. The two users can then set up a connection between them and share the file.

Digital Clock *(Aug-Nov 2011)*

Technology used: coding in BSV (BlueSpec Verilog), FPGA (Field Programmable Gate Array) architecture

- Designed a digital clock in BSV with four modes *HH:MM, MM:SS, stopwatch, alarm modes* and ran on FPGA architecture
- User could set time, alarm and use clock as stop watch

OTHER PROJECTS

Economic Analysis of Indoor Air pollution and Respiratory Health *(March-April 2012)*

- Analyzed impact of indoor air pollution caused by use of solid fuel on the health of people exposed to it.
- Economic analysis of programs and policies undertaken by Indian Government and various NGOs and cause of their failure
- Proposed new policies that government of India could undertake and also provided the technical solution to decrease the Indoor air pollution and thereby reducing the health impact

Cricbot (Robotics Club, IIT Kanpur) *(Dec 2011)*

Technology used: coding using OpenCV, Image Processing

- Autonomous robot that works on the principle of Image Processing
- The robot had to collect the balls which was rolled down from the ramp and then deposit them into the collection pit in minimum possible time and play a one-on-one cricket match against the opponent robot

Rover Bot (Robotic Club, IIT Kanpur) *(May-July 2011)*

- Developed a system design for a difficult **regional exploration** rover for demonstration of locomotion capabilities, payload accommodation, and control. This is somewhat similar to lunar rover
- Efficient locomotion system capable of moving on rough terrains, steps, cylindrical objects, and slope up to 50 degree
- An article for the same was **published** in a reputed newspaper describing about the robot's technologies.

Flying Dragon Model (Introduction to manufacturing process, IIT Kanpur) *(Jan-April 2012)*

Technology used: Sheet metal, Welding, Casting

- Made a Flying dragon model using Welding, Casting and Sheet Metal work which was attached to base at only single point
- Got **Best Project Award** among 60 other projects.

Paddle Boat (Introduction to Manufacturing Process, IIT Kanpur) *(Jan-April 2012)*

Technology used: Milling, Drilling, Lathe Machine

- Made a mechanical model of Da-Vinci Paddle Boat which supports conventional paddling by legs
- Multiple workers can provide their efforts on a central rod which prevents the problem of **synchronizing efforts** in usual paddling. Got **Second Best Project Award** among 65 other projects.

RELEVANT COURSES

- Operating Systems
- Computer Networks
- Data Structures & Algorithms
- Algorithms II
- Principles of Database Management
- Principles of Programming Language
- Special topics in Computer Science
- Introduction to electronics
- Introduction to Software Engineering*
- Compiler Design
- Artificial Intelligence Programming
- Discrete Mathematics
- Theory of Computation
- Intro. to Computer Organization
- Programming Tools & Techniques
- Economic Analysis of Laws
- Introduction to Indian Society
- Randomized Algorithms*
- Introduction to Mathematical Logic
- Probability and Statistics
- Multivariable Calculus
- Complex Analysis & Linear Algebra
- Fourier Analysis & Differential Equations
- Introduction to economics
- Fundamental of computing
- Applied Game Theory*

*to be completed in July-Nov 2013

TECHNICAL SKILLS

Programming Languages	C, C++, Java, Python, Oz, small talk, Assembly Language ,Bluespec Verilog
Platforms	Windows, Linux
Tools	LaTeX, Beamer, Yacc, Make, Shell, awk, GNU Octave, SQL, Gdb, MATLAB, Autocad

POSITIONS OF RESPONSIBILITY

- **Academic Mentor, Counseling Service** IIT Kanpur for academic year 2011-12
  - Taught ESC101(C language) and PHY103 (Electrodynamics) to students facing academic problems
- **Student Guide, Counseling Service**, IIT Kanpur for academic year 2011-12
  - Mentored 6 freshmen students and assisted them in getting familiar to the college environment
  - Responsible for their overall adjustment and performance in first year
  - Assisted in the successful organization of the orientation programme for the benefit of around 815 students in IIT Kanpur
- **Secretary, Robotics Club**, IIT Kanpur for academic year 2011-12
  - **Guided students** in robotics events in Techkriti, annual inter-collegiate technical festival of IITK and Takneek, the Intra-college

Technical Festival of IITK. Responsible for scheduling and smooth conduction of robotics **lectures and workshops**
    - Organized competitions under Takneek and Techkriti, and projects for students over the summers
- **Computer Centre Secretary**, Hall 2 for academic year 2011-12
    - Responsible for installation of new softwares and proper working of the hardware
- **Takneek Pool Co-ordinator,** Takneek'12
    - Organized scientific and technological events from *Rajput pool,* consisting of 3 hostels and ensured healthy participation. Overall efforts lead to **first position** among other pools breaking all the records. We stood first in around 75% of the event scoring 1100 points. Previous highest was around 750 only.

Extra-Curricular Activities/Interests

- **Robotics**
    - Made an autonomous line following robot in Takneek'10
    - Participated in Kshitij'11 (IIT Kharagpur technical fest).
    - Made an All-Terrain Vehicle in Summer Camp'11 organized by Robotics club IIT Kanpur
    - Participated in Wild Soccer in Takneek'11
    - Participated in Techfest'12(IIT Bombay Technical Fest).
- **Business**
    - Won **First prize** in Business simulation game(a virtual company of cycle was given and investments were affected like original market) held in IIT Kanpur
    - Going to Represent IIT Kanpur in IIM Ahmedabad in simulation games
- **Social Services**
- Awarded **Certificate of Special Effort** in Mass Awareness Campaign against AIDS and Cancer organized by Caring Souls Foundation to alleviate the sufferings of the Needy Cancer Patients
- Member of Anti ragging committee for my hall
- **Programming:**
    - Stood second in Xceed Kanpur '13 organised by Kurukshetra'13 organized by Anna university
- **Others:**
    - Scored highest marks in chemistry and Mathematics in the CBSE examination of class 12 in my school
    - Received an excellence award in field of Academics and declared Scholar for the academic year 2008-09
    - Awarded first position in elocution competition held in school in academic year 2005-2006
    - Awarded first position in Debate competition held in school in academic year 2006-2007 and second position in year 2007-2008
    - Awarded Second position in News reading competition held in school in academic year 2007-2008
    - Awarded first position in National Anthem competition held in school in academic year 2007-2008
    - Awarded second position in tableau competition held in school in academic year 2007-2008
    - Awarded second position in skit competition held in school in academic year 2005-2006
    - Football team got gold medal in intra school football match.