Final Thoughts and Conclusions

(Take away's from this course!)

Sharma Chakravarthy
Information **T**echnology **Lab**oratory (IT Lab)
Computer Science and Engineering Department
The University of Texas at Arlington, Arlington, TX 76019
Email: sharma@cse.uta.edu
URL: http://itlab.uta.edu/sharma

Other Courses I Teach + Research

- I teach cse 3330/5330 (Intro to Databases)
- I teach cse 5334 (Data Mining)
- I teach cse 4331/5331 (database II: systems aspects)
- I also teach DASC 5300/CSE 5300
- I also teach cse 6331, 6332 (advanced topics in databases), cse 6339 (seminar course), and other courses
- I have taught CSE 1325 (Java)
- If you would like to do more challenging projects in Big Data Analytics, Mining, Machine Learning, video situation analysis, please come and talk to me (MS thesis, projects, …)
- Most of our research prototypes are implemented in Java (MavStream, Sentinel, InfoSift, …), Python, C, C++
  - We have 50,000+ lines of code in Java
- I will be happy to involve you in one of the research projects

Databases is a vast area

- Object-Oriented DBMSs
- Object-Relational DBMSs
- Complex event processing (CEP) or Active DBMSSs (I offer)
- Data Stream Management Systems (DSMS) (I offer as 5339)
- Cloud computing and DBMSs (I offer as 6331)
- XML Support in DBMSs
- Distributed and parallel DBMSs (have offered in the past; has ties with block chain)
- Multi-media DBMSs
- Data Mining (I offer as cse 5334)
- Data Warehousing (have offered in the past)
- Spatial and temporal databases
- Time series or sequence analysis
- Information security

Data Science/Analytics is even Vaster

- Combines work from
  - DBMSs
  - Mining approaches
  - Machine learning algorithms
  - Knowledge discovery
  - Wider analysis of objectives
  - 4V's (volume, variety, velocity and veracity)
  - Information integration or fusion
  - Multiple data models for representing 4V data

The devil is in the detail

- Hope this course has provided some idea of the internals of a RDBMS and how parts of it are implemented
  - More importantly, each topic/technique is useful in its own right for applications beyond RDBMS
  - Hope you can build further using this as a foundation!
- CSE 3330/5330 covers how to go from a real-world problem to a database design and implementation
- This course takes it from there and covers important modules and their details
  - Storage and indexing
  - Concurrency control and recovery
  - Query optimization
  - Newer models and NoSQL systems
  - Cloud computing and Map/Reduce
- Interaction between the components of a DBMS

Feedback

- Send email to: sharmac@cse.uta.edu or on Polly quiz today
- What works and what does not work. More home works? More quizzes? How to make students practice more?
- How do I convey the seriousness about starting the project early
- What component of the course is hard to grasp? What can I do about it?
- How can I make the project more interesting?
- How do I increase class participation?
- Any feedback on the TA, project, and how to improve his/her role for doing better in this course

Thank You !!!

I have enjoyed teaching the course. Hope you have too!

For more information visit:

http://itlab.uta.edu