
COLLEGE OF ENGINEERING, UIC
CS 594, Vector Databases, 4 credits

I. Instructor & Course Details

Instructor Name: Abolfazl Asudeh

<https://asudeh.github.io/>

Email address: asudeh@uic.edu

Drop-In Office Hours (in-person): Mondays, 1:00 PM to 3:00 PM

Drop-In Hours location (*office number*): CDRLC 5452

Course Modality and Schedule (*UIC Syllabus Policy and HLC PRIORITY SECTION*)

(Public) Course Page: <https://asudeh.github.io/teaching/archive/cs594fall26/>

This course is taught **ON CAMPUS**.

- DAYS and TIMES: Mondays and Wednesdays, 9:30 AM-10:45 AM
- LOCATION: CDRLC 1409

II. Course Information

Catalog Course Description and Prerequisite/Corequisite Statement

CS 594: Vector Databases is a graduate special-topics course on vector databases. This section studies the foundations, algorithms, and systems used to store, index, query, and manage high-dimensional vector representations. The course connects database systems, information retrieval, machine learning, and algorithms.

Topics include vector representations and embedding models, similarity measures, exact and approximate nearest-neighbor search, locality-sensitive hashing, product quantization, graph-based and disk-based ANN indexes, vector-database architectures and query models, filtered and hybrid search, query planning, retrieval-augmented generation (RAG), and related data models such as knowledge graphs.

Credit: 4 graduate hours.

Prerequisite: CS 401, 480; Consent of the instructor, consistent with the UIC catalog for CS 594. Background in algorithms and data structures, basic database systems, and familiarity with machine learning or information retrieval is needed.

1. CS 401 (Algorithms I) or equivalent background in algorithm analysis and data structures.

2. CS 480: Database systems
3. Recommended: prior exposure to databases, machine learning, data mining, or information retrieval.

A high-level outline of topics:

- Introduction to vector databases and vector similarity search
- Vector representations, embedding models, and retrieval evaluation
- Exact and approximate nearest-neighbor search in high dimensions
- Locality-sensitive hashing and partition-based ANN methods
- Inverted-file indexes, vector quantization, and product quantization
- Graph-based ANN search, including HNSW and related graph indexes
- Disk-based and memory-efficient ANN search
- Vector-database architecture, storage, indexing, and query execution
- Filtered, constrained, and hybrid vector search; query planning and optimization
- Retrieval-augmented generation (RAG) over vector databases
- Knowledge graphs, hybrid semantic retrieval, and open research problems in vector data management

Required and Recommended Course Materials

- There is no required textbook. Readings will consist of research papers, surveys, and selected book chapters listed on the public course page. Core references include Pan, Wang, and Li, "Survey of Vector Database Management Systems"; Han, Liu, and Wang, "A Comprehensive Survey on Vector Database"; and the research papers assigned for individual indexing and retrieval methods.

Respect for Copyright

Please protect the copyright integrity of all course materials and content. Please do not upload course materials not created by you onto third-party websites or share content with anyone not enrolled in our course.

III. COURSE POLICIES & CLASSROOM EXPECTATIONS

Grading Policy and Point Breakdown

The course is project- and research-oriented; there is no final exam. Coursework emphasizes technical understanding, implementation, experimental evaluation, research reading, and communication. Unless an assignment explicitly permits collaboration, submitted work must be the student's own.

Scribe Notes & Paper Discussion 20%

- Students will prepare clear technical scribe notes for assigned course material and critically engage with the associated research papers. The notes should explain the problem, core technical ideas, assumptions, guarantees or design tradeoffs, and limitations rather than only summarizing the lecture.
- Scribe-note assignments and submission format will be announced on the course page or LMS. Proper attribution and citation of all external material are required.

Programming Project 30%

- Students will implement core components of a vector database and build a RAG system on top of the implementation. Detailed milestones, interfaces, datasets, and evaluation criteria will be provided separately.

Research Project: 40%

- Technical Merit
- Presentation
- Report

Active Class Participation 10%

Grading Policy

Course grades **will be curved as necessary**, based on each student's total numeric score for all coursework at the end of the semester. In general, the marking scheme for this class will be as follows:

A = 90-100; B = 80-89; C = 70-79; D = 60-70; F = Below 60.

These standards may be adjusted for certain exams or homework. These adjustments will be announced in class as the exam/homework is handed back.

Any request for reconsideration of any grading on coursework **must be submitted within one week of when it is returned**. Any coursework submitted for reconsideration may be regarded in its entirety, which could result in a lower score if warranted.

Depending on class performance, the scale may be adjusted to compensate (e.g. 89% may become an A) but at the very least, the cutoffs listed above are guaranteed (e.g. a 91% will not

become a B). You can use this straight grading scale as an indicator of your minimum grade in the course at any time during the course. You should keep track of your own points so that at any time during the semester you may calculate your minimum grade based on the total number of points possible at that particular time.

Policy for Missed or Late Work

Late work: Late submissions will not, in general, be accepted. They will never be accepted if the student has not made special arrangements with me at least one day before the deadline. If late work is accepted, it may be subject to a score reduction.

Incompletes: The UIC Undergraduate catalog states that in addition to needing excellent justification for an incomplete, a student must also have been “making satisfactory progress” in the course.

Statute of limitations: No grading questions or complaints, no matter how justified, will be listened to one week after the item in question has been returned.

Cheating: Cheating will not be tolerated. All work you submitted must be entirely your own. Any suspicious similarities between students' work (this includes homework and exams) will be recorded and brought to the attention of the Dean. The MINIMUM penalty for any student found cheating will be to receive a 0 for the item in question, and dropping your final course grade one letter. The MAXIMUM penalty will be expulsion from the University.

Classroom Conduct: Classroom discussions and questions are a valuable part of the learning process and are encouraged. However, students who repeatedly talk among themselves disrupting the class lecture will be asked to leave.

Attendance / Participation Policy

Please email me if you face an unexpected situation that may impede your attendance, participation in required class and exam sessions, or timely completion of assignments.

Other Course Policies: *(UIC Syllabus Policy and HLC PRIORITY SECTION)*

Academic Integrity

UIC is an academic community committed to providing an environment in which research, learning, and scholarship can flourish and in which all endeavors are guided by academic and professional integrity. In this community, all members including faculty, administrators, staff, and students alike share the responsibility to uphold the highest standards of academic honesty and quality of academic work so that such a collegial and productive environment exists.

As a student and member of the UIC community, you are expected to adhere to the Community Standards of integrity, accountability, and respect in all of your academic endeavors. When accusations of academic dishonesty occur, the Office of the Dean of Students investigates and adjudicates suspected violations of this student code. Unacceptable behavior includes cheating, unauthorized collaboration, fabrication or falsification, plagiarism, multiple submissions without instructor permission, using unauthorized study aids, coercion regarding grading or

evaluation of coursework, and facilitating academic misconduct. Please review the UIC Student Disciplinary Policy for additional information about the process by which instances of academic misconduct are handled towards the goal of developing responsible student behavior.

By submitting your assignments for grading you acknowledge these terms, you declare that your work is solely your own, and you promise that, unless authorized by the instructor or proctor, you have not communicated with anyone in any way during an exam or other online assessment, and that submitted work is your own unless the assignment explicitly is intended to be done in a group. Use of Generative AI and Large Language Models is permitted only when explicitly authorized for a specific assignment or project. Otherwise, such use is not allowed.

We use an automatic cheating-verification program that is capable of detecting partial logical similarities of answers. Don't even take the risk! Do NOT use or post to Chegg or any similar Website.

Plagiarism is a serious matter and will be treated as such. You may be penalized by failing the course. Any student caught will have a grade of zero on the assignment, will have a drop in their letter grade at the end of the semester, and will be reported to the Dean of Students. This applies to each offense. Details on this are given on the Academic Integrity page (link: <https://dos.uic.edu/community-standards/academic-integrity/>) and you can view the Student Disciplinary Policy at this link: <https://dos.uic.edu/wp-content/uploads/sites/262/2018/10/DOS-Student-Disciplinary-Policy-2018-2019-FINAL.pdf>).

We believe that each and every one of you is able to learn the material and complete assignments on your own. If you are struggling, please reach out to the instructional staff and ask for help rather than submitting work that you did not complete yourself. It is better to earn a lower score on an assignment than to incur an Academic Integrity violation.

Email Expectations

Students are responsible for all information instructors send to your **UIC email and Canvas accounts**. Faculty messages should be regularly monitored and read in a timely fashion. All critical announcements, changes to assignments, etc. will be announced with **Canvas announcement**. We are assuming that you check your email regularly, at the very least once every 24 hours

IV. COURSE SCHEDULE

Tentative Weekly Schedule of Class Topics, Readings, and Course Milestones

Date	Topic	Reading	Notes
Week 1 M. 8/24	Course Introduction; What is a Vector Database?	Pan, Wang & Li: VDBMS Survey; Han, Liu & Wang: Vector Database Survey	
W. 8/26	Vector Representations and Similarity Measures	Stanford IR Book, Ch. 6; Pan, Wang & Li survey	
Week 2 M. 8/31	Embedding Models: Text, Image, and Multimodal Embeddings	Jurafsky & Martin, SLP Ch. 5; word2vec; GloVe	
W. 9/2	Embedding Quality, Dimensionality, and Retrieval Evaluation	Stanford IR Book, Ch. 8	
Week 3 M. 9/7	Labor Day Holiday -- No Class		No Class
W. 9/9	Exact Nearest-Neighbor Search and the Curse of Dimensionality	Han, Liu & Wang survey	
Week 4 M. 9/14	Approximate Nearest-Neighbor Search: Problem Formulation and Guarantees	Han, Liu & Wang survey	
W. 9/16	Locality-Sensitive Hashing (LSH)	Gionis, Indyk & Motwani: Similarity Search in High Dimensions via Hashing	
Week 5 M. 9/21	Tree- and Partition-Based ANN Indexes	Guttman: R-trees; Han, Liu & Wang survey	
W. 9/23	Inverted File Indexes (IVF) and Clustering-Based Search	Pan, Wang & Li survey	
Week 6 M. 9/28	Vector Quantization and Product Quantization	Jegou et al.: Product Quantization; Ge et al.: Optimized Product Quantization	
W. 9/30	Graph-Based ANN Search	Malkov & Yashunin: HNSW; Fu et al.: NSG; Dehghankar & Asudeh: HENN	
Week 7 M. 10/5	HNSW: Construction, Search, and Design Tradeoffs	Malkov & Yashunin: HNSW; k-NN graph indexing optimization	
W. 10/7	Disk-Based and Memory-Efficient ANN Search	DiskANN; Graph-based ANN survey	
Week 8 M. 10/12	Vector Database Architecture: Storage, Indexing, and Query Execution	Pan, Wang & Li survey; VDBMS tutorial	
W. 10/14	Vector Database Systems and Design Tradeoffs	Pan, Wang & Li survey; Han, Liu & Wang survey	
Week 9 M. 10/19	Vector Query Models and Top-k Similarity Search	Stanford IR Book, Chs. 6, 11, 12; Pan, Wang & Li survey	
W. 10/21	Filtered and Constrained ANN Search	Filtered-DiskANN; ACORN; ANN search with attribute constraints	
Week 10 M. 10/26	Hybrid Search: Combining Structured Filters and Vector Similarity	Pan, Wang & Li survey	
W. 10/28	Query Planning and Optimization for Vector Search	Pan, Wang & Li survey; VDBMS tutorial	
Week 11 M. 11/2	Retrieval-Augmented Generation (RAG): Architecture and Retrieval Pipeline	Gao et al.: Retrieval-Augmented Generation for Large Language Models: A Survey	
W. 11/4	RAG Retrieval: Chunking, Embeddings, Reranking, and Context Selection	Gao et al.: RAG Survey	
Week 12 M. 11/9	Building RAG Systems on Vector Databases	Gao et al.: RAG Survey	
W. 11/11	Evaluating Vector Retrieval and RAG Systems	Assigned papers / course page	
Week 13 M. 11/16	Knowledge Graphs and Graph-Based Retrieval	Assigned papers / course page	

Date	Topic	Reading	Notes
W. 11/18	Knowledge Graphs + Vector Databases; Hybrid Semantic Retrieval	Assigned papers / course page	
Week 14 M. 11/23	Research Directions in Vector Databases and Unstructured Data Processing	Pan, Wang & Li survey; Han, Liu & Wang survey	
W. 11/25	Student Wellness Day -- No Class		No Class
Week 15 M. 11/30	Student Project Presentations / Research Discussions		Project Presentations
W. 12/2	Student Project Presentations / Course Wrap-up		Project Presentations

Disclaimer

This syllabus is intended to give the student guidance on what may be covered during the semester and will be followed as closely as possible. However, as the instructor, I reserve the right to modify, supplement, and make changes as course needs arise. I will communicate such changes in advance through in-class announcements and in writing via Blackboard Announcements.

V. ACCOMMODATIONS

Disability Accommodation Procedures

UIC is committed to full inclusion and participation of people with disabilities in all aspects of university life. If you face or anticipate disability-related barriers while at UIC, please connect with the Disability Resource Center (DRC) at drc.uic.edu, via email at drc@uic.edu, or call (312) 413-2183 to create a plan for reasonable accommodations. To receive accommodations, you will need to disclose the disability to the DRC, complete an interactive registration process with the DRC, and provide me with a Letter of Accommodation (LOA). Upon receipt of an LOA, I will gladly work with you and the DRC to implement approved accommodations.

Religious Accommodations

Following [campus policy](#), if you wish to observe religious holidays, you must notify me by the tenth day of the semester. If the religious holiday is observed on or before the tenth day of the semester, you must notify me at least five days before you will be absent. Please submit [this form](#) by email with the subject heading: **"YOUR NAME: Requesting Religious Accommodation."**

VI. CLASSROOM ENVIRONMENT

Inclusive Community

UIC values diversity and inclusion. Regardless of age, disability, ethnicity, race, gender, gender identity, sexual orientation, socioeconomic status, geographic background, religion, political ideology, language, or culture, we expect all members of this class to contribute to a respectful, welcoming, and inclusive environment for every other member of our class. If aspects of this course result in barriers to your inclusion, engagement, accurate assessment, or achievement, please notify me as soon as possible.

Name and Pronoun Use

If your name does not match the name on my class roster, please let me know as soon as possible. My pronouns are [he/him](#). I welcome your pronouns if you would like to share them with me. For more information about pronouns, see this page: <https://www.mypronouns.org/what-and-why>.

Community Agreement/Classroom Conduct Policy

- Be present **by turning off cell phones** and removing yourself from other distractions.
- Be respectful of the learning space and community. For example, **no side conversations** or unnecessary disruptions.
- Use preferred names and gender pronouns.
- **Assume goodwill in all interactions**, even in disagreement.
- Facilitate dialogue and value the free and safe exchange of ideas.
- Try not to make assumptions, have an open mind, seek to understand, and not judge.
- Approach discussion, challenges, and different perspectives as an opportunity to “**think out loud**,” learn something new, and understand the concepts or experiences that guide other people’s thinking.
- Debate the concepts, not the person.
- Be gracious and open to change when your ideas, arguments, or positions do not work or are proven wrong.
- Be willing to work together and **share helpful study strategies**.
- **Be mindful of one another’s privacy**, and do not invite outsiders into our classroom.

Content Notices and Trigger Warnings

Our classroom provides an open space for a critical and civil exchange of ideas, inclusive of a variety of perspectives and positions. Some readings and other content may expose you to ideas, subjects, or views that may challenge you, cause you discomfort, or recall past negative experiences or traumas. I intend to discuss all subjects with dignity and humanity, as well as with rigor and respect for scholarly inquiry. If you would like me to be aware of a specific topic of concern, please email or visit my Student Drop-In Hours.

VII. RESOURCES: Academic Success, Wellness, and Safety

We all need the help and the support of our UIC community. Please visit my **office hours** for course consultation and other academic or research topics. For additional assistance, please contact your assigned college advisor and visit the support services available to all UIC students.

Academic Success

- [UIC Tutoring Resources](#)
- College of Engineering [tutoring program](#)
- [Equity and Inclusion in Engineering Program](#)
- [UIC Library](#) and [UIC Library Research Guides](#).
- [Offices](#) supporting the UIC Undergraduate Experience and Academic Programs.
- [Student Guide for Information Technology](#)

Wellness

- **Counseling Services:** You may seek free and confidential services from the Counseling Center at <https://counseling.uic.edu/>.
- Access [U&I Care Program](#) for assistance with personal hardships.
- **Campus Advocacy Network:** Under Title IX, you have the right to an education free from any form of gender-based violence or discrimination. To make a report, email TitleIX@uic.edu. For more information or confidential victim services and advocacy, visit UIC's Campus Advocacy Network at <http://can.uic.edu/>.

Safety

- [UIC Safe App](#)—PLEASE DOWNLOAD FOR YOUR SAFETY!
- [UIC Safety Tips and Resources](#)
- [Night Ride](#)
- [Emergency Communications](#): By dialing 5-5555 from a campus phone, you can summon the Police or Fire for any on-campus emergency. You may also set up the complete number, (312) 355-5555, on speed dial on your cell phone.

Syllabus Revision

The standards and requirements set forth in this syllabus may be modified at any time by the course instructor. Notice of such changes will be by Canvas announcement or email notice.

Disclaimer

This syllabus is intended to give the student guidance in what may be covered during the semester and will be followed as closely as possible. All differences between information in this

syllabus and the official UIC academic calendar should be resolved in favor of the calendar. However, as instructors, we reserve the right to modify, supplement, and make changes as course needs arise. We will communicate such changes in advance as Blackboard announcements.