With all the courses we're able to take and the I N T E R D I S C I P L I N A R Y nature of biomedical informatics it can be difficult to find courses that match your interests, or know which are worth your time. To that end, this document should help students find courses taken by other students in the department and get a general feel for the material covered. So, for the greybeards of the department (essentially after your first semester you're a greybeard because academia ages you something fierce) go ahead and follow the format below. **Be sure to include information about the focus of the class, how material is presented, workload, and whether or not you got out of it what you were looking for and place courses under the requirements they fulfill.** If a course fulfills multiple requirements put it under one requirement, and mention the other requirements it fills in your review. For students looking for courses to take, feel free to peruse the ancient wisdom contained within this tome, and even reach out to the authors of said wisdom.

Current list of courses and requirements

Please follow this template:
1. Course Name, Course #
    a. Professor who taught the course. Your review. -Yourname, Semester your took the course
    b. Professor who taught the course. Another student's review. -Anotherstudentsname, Semester they took the course

Table of Contents:

Domain Courses

Clinical

1. Acculturation to Medicine and Clinical Informatics, BINF G4011
   a. Professor: Sivan Kinberg. I'd say this is a can't miss course of anyone on the clinical track who doesn't have clinical experience of their own (or even if they do). Sivan does a really nice job, and shadowing in the hospital is invaluable. The workload is manageable, with some weekly readings for discussion and a final book report. -Elliot, Spring 2017
   b. Professor: Sivan Kinberg. Great class! You learn a lot (especially for clinical students with no clinical background) without a ton of work (typically weekly readings and a few low-lift projects throughout the semester). The class really focuses on teaching you medical "lingo" and you get really cool shadowing opportunities! - Apara, Spring 2023
2. Data Science for Mobile Health, BINF G5001
   a. Professor: Noemie Elhadad and Chintan Patel. I'm not sure how often this course will be taught again, but I really enjoyed it for the hands on experience and interesting paper discussions. -Elliot, Spring 2018
3. Defining, Evaluating and Improving Quality in Health Care, BINF G5000
   a. Professor: Rimma Perotte. For anyone in the clinical track (or those interested in a low stakes intro to evaluating/improving care), this class is a must. Course readings are optional, the class is held once a week for 2 hours, and the only requirements are (minimal) participation, 1-2 pg weekly responses based on lecture (not readings), and a 10-pg final paper that is essentially an extended version of the weekly response on a single topic. Most speakers are very interesting, dynamic, and eager to engage in discussion. Rimma is awesome and supplements guest speaker presentations with "just in time" context and background that is pared down to what is actually helpful/relevant. -Harry, Spring 2021
   b. Professor: Rimma Perotte. I agree with everything Harry said! As someone with clinical experience, I found this class incredibly useful to give the "things we do for no reason" an actual reason (albeit, not a medical one). - Lauren, Spring 2021
   c. Professor: Rimma Perotte, David Vawdrey. This class is a fantastic, multi-faceted look at what "quality" means in a variety of healthcare contexts. The format remained the same as spring 2021, but we had the added bonus of David Vawdrey (from Geisinger) as another professor of the class. In addition to the above, I'll add that Rimma is super keen on collecting feedback and really tailoring the invited speaker lists based on mini-comment cards that we filled out

each week; it was a great opportunity to provide input on what topics felt highly relevant or which might need to be phased out and replaced with other lectures (e.g. COVID-specific content). -Courtney, Spring 2022

4. Introduction to Randomized Clinical Trials, BIST P8140
    a. Prof: John (Seamus) Thompson. I took this course several years ago for my MPH, but from what I remember it was a decently valuable course. It covers important elements of study design and analysis for clinical trials. Developing an entire Statistical Analysis Plan for a hypothetical study as the final project was interesting and probably helped to develop some valuable skills for putting together research proposals/planning studies. It was a relatively slow-paced course, very easy to keep up with the content, and most of the real work was concentrated later in the semester with the final project (my advice: don't procrastinate!). – Kristen King, Spring 2017

5. Epi Modeling for Infectious Diseases, EPID P8477
    a. Prof: Wan Yang. An interesting class, kind of a niche topic, but the professor was super nice and enthusiastic and I think her COVID-19 modeling was really important to the NYC DOH early in the pandemic! If you haven't taken calculus recently the equations during lecture can look kind of scary, but most of the code is given to you (in R) and assignments focus more on knowing what parameters to plug into the different models in various scenarios and how to interpret the output. A good amount of class time is reserved to work on the final group project so the workload is very reasonable in my opinion. – Kristen King, Spring 2018

6. Pharmaceutical Statistics, BIST P8144
    a. Prof: Naitee Ting. Honestly would not recommend this one. The content is very specific to pharmaceutical clinical trials (Phases I-III) and was targeted as if every student was planning a future career as a biostatistician at a pharma company working to get new drugs approved. When I took it the class was not very structured with lots of old school whiteboard writing during lectures, repetition in content from class to class, and I had trouble staying awake/engaged during lectures. I think other classes cover similar content in ways that would be more relevant to the field of biomedical informatics (BIST P8140 Intro to RCTs, BINF G6002 Research Methods, maybe the intro Biostatistics methods class under Quant). But it's easy/not too much work if you just need a filler class one semester. – Kristen King, Spring 2018

Translational

1. Mechanisms of Human Disease, G6003
    a. Professor: Ronald Liem. Overall I really enjoyed this class. This course has three modules where each module focuses on a different disease (Sickle Cell, Alzheimer's and Atherosclerosis the first semester) and each lecture is taught by a different professor. Workload is light the first two

modules with no readings and a take home exam following each module. Exams are open notes/papers and you are given a week to complete them. They are on the creative side instead of just regurgitating facts mentioned in class. Familiarity with wet lab techniques is helpful since there is often a "design an experiment" portion of the exam. The third module you have to read (generally) wet-lab based papers before each class and post a question. A student then presents that paper during class and answers everyone's questions (ex. of an assigned paper [10.1096/fj.201701381)](10.1096/fj.201701381). This presentation counts as your third module exam grade. Pre-reqs: I've never taken biochem before and was fine, basic bio knowledge is helpful. - Katie Brown, Fall 2019

2. Translational Bioinformatics
   a. Professor: Nick Tatonetti. The class is essentially a grab-bag of different topics related to bioinformatics, and stuff directly from Nick's research. He covers basic ML stuff, bioinformatics methods like GSEA, and more clinical topics like ontology structure. When I took it, there was one lecture per week and one lab session, which were both taught in person on the morningside campus. The class covers many different interesting topics, but didn't have a clear progression in my opinion. The assignments were all 'optional,' and you'll do fine on the final project if you put in a reasonable amount of effort. FYI, this course is typically offered every other fall semester, on odd numbered years, so take it when it's being offered. Jonah Einson, Fall 2017

Biological

1. Computational Genomics, CBMF 4761 W
   a. Professor: Itsik Pe'er. I am by no means a genetics guy. I've worked a bit with them, and I've heard they even determine a good chunk of my future, but this class felt like a good balance between algorithms I've seen elsewhere in my ML classes, and genomics. It's an absolute ton of work, and you'll spend a lot of time trying to decipher what the questions are even asking, but once you do you'll feel like you can handle anything… until the next homework comes out. The lectures and slides are very informative and I think Itsik provides a good amount of information that prepares you for the assignments, but allows you to stretch too. If you take this class be sure to rope a few other students in with you as taking this class with friends helps a lot.-Oliver Bear Don't Walk IV, Fall 2018
   b. I am more of a genetics guy, and I really enjoyed this class. A lot of Bioinformatics courses teach generally the same thing, but Computational Genomics pushes you harder to learn the nitty gritty of how commonly used algorithms actually work. Itsik is a good lecturer, but his assignments can be quite

challenging. Overall, would definitely recommend for anyone on the Bio track. - Jonah Einson, Spring 2018

    c. This class has some heavy(but interesting) CS and stats content, focused on the theoretical basis for many common tasks in genomics - sequence alignment, genome assembly, variant calling, RNA-seq quantification etc. The problem sets were pretty hard and took a lot of work. Would definitely recommend taking it with a group of people so you can work together on projects (sys bio phd students usually take it as well so could work with them thats what I did) - Vinay Swamy, spring 2022

    d. This class is HARD. Like, quite hard. And a lot of work! But, it is graded very reasonably (most folks end up with a decent grade if they do the work) and the content is quite relevant if you are interested in anything genomics. From sequence alignment, mapping, and genome assembly, to single-cell dimensionality reduction, hidden markov models, ML, phylogenetic trees etc. I highly recommend engaging in class and with the material, because all of the concepts help you if you use bowtie2, samtools, tSNE, assembly, variant callers etc. in your workflow. Definitely collaborate with others, and start thinking about the final project really early on. - Aziz Zafar, Spring 2024

2. Deep Sequencing
    a. Professors: Peter Sims, Yufeng Shen, Choalin Zheng. Great course and very comprehensive introduction to high throughput sequencing and its statistical analysis. There's a project component that is a great learning experience. Essential for understanding one of the pillars of systems biology. - Nick Giangreco

3. Statistical Genetic Modeling
    a. Professors: Shuang Wang. This is a required course for biostatistics majors, but was helpful for me, who uses these concepts a lot in my own research. If you're doing anything pop-gen related, definitely a good course to take. Shuang is a decent lecturer, and really makes sure you understand the material. The exams were designed to push you to your limit… - Jonah Einson, Spring 2019

4. Machine Learning for Functional Genomics
    a. Professors: David Knowles. This can either count as a Bioinformatics domain or Quant course. I **strongly** recommend this course for anyone interested in some of the math/theory behind machine learning and its direct relevance to the application in functional genomics/transcriptomics. I am someone who came in very familiar with Applied ML in scikit-learn and Torch. This course did a wonderful job at towing the line between theory and application for my background. A broad range of topics are covered, and covered well – from basics of Ridge and Lasso to Hidden Markov Models to Autoencoders. Knowles is a great instructor, and the class is very hybrid-friendly. There are 3-4 assignments, which are a decent amount of work, but overall the class is not very heavy in its workload. Just attend/watch lectures, learn, and implement semi-applied machine

learning in the assignments/homeworks. - Aziz Zafar, Fall 2023. P.S. This is one of the best courses I've **ever** taken.

5. Statistical Machine Learning for Genomics, BMCS 4480
    a. Professors: Elham Azizi. I think this course is intended to be a parallel to Knowles's course (reviewed above) with a focus on single-cell data. The early assignments are half basic scanpy and half problem sets, with the remaining assignments being journal club-type presentations and a partner project. Most of the class found the problem sets difficult, but they don't count for a lot of your grade and aren't really important for succeeding later in the course. That said, if you don't have a level of mathematical maturity equivalent to what you'd get in the CS department's ML course, try to partner with someone who does. You don't really get practice with anything beyond clustering/regression and the expectations for the project are relatively high, though Elham/the TAs were helpful for talking through it. I felt like I understood the state of single-cell methods research well after taking the class, but a lot of the course content wasn't presented very clearly and it's about as demanding as you'd expect for a CS class. I would recommend Machine Learning for Functional Genomics or Computational Genomics if you're more broadly interested in computational methods for genomics, but if you want some coursework related to single-cell or have a project idea you're excited about then you might get something out of it. - Chris Allsman, Spring 2024

6. Topics in Mathematical Genomics - Applications, CMBS5301GR
    a. Professors: Tal Korem, Raul Rabadan. This is the "Intro" equivalent for the Bioinformatics track students. One lecture a week, followed by a journal club - very little workload, very reasonably graded/assessed, and very educational and informative. Topics range from Cancer Genomics and Virology (Raul's portion) to Microbiome analysis (Tal's portion) to genomics and protein modeling (Mohammad Al Quraishi, Anum Glasgow, and Yufeng Shen) and even some EHR (Noemie Elhadad). Topics and presenters may change from year to year, but the instructors do a good job making sure students get a taste of different projects and directions within Computational Biology/Genomics. Highly recommended. - Aziz Zafar, Spring 2024

Public Health

Quantitative(Quant)

1. Probabalistic Machine Learning (Formerly: Foundations of Graphical Models), STCS 6701 GR
    a. Professor: Dave Blei. Easily one of the best classes I've taken here. Dave Blei is hella chill about homeworks as this is a true graduate level course. However, I did

still learn a ton about probabilistic graphical modeling with a good mix of applied and theoretical problems in the homeworks. I think what made this course is the instructor. He's insanely smart when it comes to his craft, but he's one of the few professors who can also convey this knowledge in an easy to understand way while staying true to the theory. He's also just really funny (Walnut Crew where you at?). I think this course was a fair amount of work for the unit count. Also, even though I don't use this material a lot in my own research I think it helped me understand research going on in the department a lot better.-Oliver Bear Don't Walk IV, Fall 2018

b. Professor: Dave Blei. Probably the most intellectually stimulating and useful classes I've taken. Can't recommend this enough!! Dave Blei is really chill and has a level of clarity that is hard to find when explaining complicated concepts. I love how interactive this class was as well. Dave loves to learn everyone's names and interact with the class. Bear in mind, if you don't have enough background in probability theory, I would not recommend taking this course up front. Take computational methods prior to taking this course (if you don't already have the background). Workload wise it's pretty light apart from the final project which is the majority of the grade. Intellectually, this is a very heavy workload course. If you want to get deep into the weeds of how graphical models work, you have to spend significant time in reading about the theory. The readings are really fantastic references and I still refer to them in my research now. This is a can't miss for anyone interested in probabilistic modeling, deep learning, and quantitative methods! - Shreyas, Fall 2018

c. Professor: Dave Blei, inventor of the LDA (shoutout to Vivian <3). I echo most of the sentiments above. Blei is an excellent lecturer, really approachable and I found him to be just a nice professor. :) Insanely smart but actually his lectures are pretty easy to follow and he's coming out with a book which he shares in class and that helped me a lot in understanding the material. I did think this course was a ton of work (you have to read 1 paper per week) there are 3 homeworks + final project (with a paper and poster presentation). Advise taking this with some friends. I was alone. –Cindy, Fall 2023

2. Neural Networks and Deep Learning, ECBM 4040 E
    a. Professor: Mehmet Turkcan. The class is still taught in tensorflow and is very hands on – the goal is for you to understand (both mathematically and in terms of the code) how to code basic neural networks from scratch, and I would say that I got that. There's also some emphasis on how the practicalities of using deep learning (e.g. parameters to consider, evaluating your model) but the focus is building a foundation of how deep learning works. It's also pretty vision-heavy (lots of CNNs) but lessons are applicable to other systems. The homeworks are almost all coding and pretty manageable, but do require significant effort. The midterm was fair, the final project is a bit hit-or-miss depending on partners and the paper you end up choosing. – Apara, Fall 2023

    b.   Professor: Zoran Kostic. When I took this class it was taught in TensorFlow and was very applied. The homeworks had very little math, and were mostly focused around using different deep learning architectures and tuning them to get a certain performance. There was a midterm and a final project as well. The midterm was fair, but much more math heavy than the rest of the course, while the final project felt like a good way to flex your skill. I think The prof allowed students to recreate other published research papers which I haven't seen before but is a good sanity check for your work. The class assumes you have knowledge of how basic machine learning works so be sure you know the workflow of an ML project and how that can be applied to deep learning. I felt like I came out of this class with a solid grasp of how to implement DL projects. The lectures were also very good and full of information I still reference today. This class is also a good amount of work. As with any CS and ML class you're going to spend a lot of time debugging and trying to get that last bit of performance out of your model.-Oliver Bear Don't Walk IV, Fall 2017

3. Natural Language Processing, COMS 4705 W
   a. Professor: Kathy McKeown. I took an NLP and deep learning class before this, but this class surprised me by having the first half cover traditional NLP, which was very useful. I don't use those techniques now, but it's good to have a grounding in a more traditional approach. The second half is more about deep learning, and is taught in keras/python. It's more high level than TensorFlow but can still take some getting used to. The assignments can be pretty brutal as this is a CS class, but it was well worth it. -Oliver Bear Don't Walk IV, Fall 2017
   b. Professor: Kathy McKeown. Ditto Oliver's comments, above (I was in the same section). I really liked the combination overview of classical NLP techniques with deep learning NLP. Definitely some tough assignments and final exam, as seems to be the case with most CS department courses. -Elliot, Fall 2017

4. Applied Machine Learning, COMS 4995 W
   a. Professor: Andreas Mueller. I wanted to take a machine learning class just to get it on my resume. I was worried it would be too easy, but it surprised me and surpassed my expectations. I learned a ton of new stuff. Dr. Mueller is extremely practical and very well-informed. He does development for scikit-learn so he's very aware of everything new going on the machine learning space. In short, this class is super worth it; he's an awesome professor and it will give you a very good grounding in a wide range of topics. Great for people who are new to the space as well as the folks who just need a refresher. -Lisa, Spring 2019
   b. Professor: Andreas Mueller. I agree with Lisa. This was a great course for applied ML. I've recommended it to other students outside of DBMI who felt the same. Andreas is an awesome instructor! -Amanda, Spring 2019

5. Causal Inference, COMS 4995 W (Note: this course number will be shared for many different classes because these are basically wildcard courses that change every semester)

a. Professor: Adam Kelleher. This course covers a basic introduction to causal inference. The book used is Counterfactuals & Causal Inference: Methods and Principles of Social Science Research (otherwise known as Morgan & Winship). This is a pretty gentle (not super math heavy) introduction to the fundamental concepts of causal inference. Lecture was once a week in a 3 hour chunk. Adam has worked in many companies (Barcalay's when I was in the course) and knows all the directly applicable concepts of causal inference to the real world -- so he does a really good job of connecting theory to practice. I will say lectures can be very long and quite dry. I think I got more out of reading the book than lecture. Adam is a really nice guy though. Some of the conversations I had with him after class were great. Take this class if you want a gentle intro to causal inference and already have some probability background. I'd say this is a fairly low workload class. There's a midterm and a final but both are not too difficult. Be weary of the long winded, dry lectures though! - Shreyas, Fall 2018, PS: I believe there's another course on causal inference offered by David Blei in spring, haven't taken it but he's a fantastic professor.

b. Professor: Elias Bareinboim. I think this was one of the hardest classes I've ever taken, but was also very rewarding. There were around 5 homeworks throughout the semester along with a midterm and final exam. I would say each homework took me around 20 hours to complete. The office hours are super useful, I would come prepared with specific questions. Despite how hard this class was I do think it was beneficial in introducing me to a whole new field of study, and influencing how I think about certain problems. I wouldn't recommend taking another demanding class at the same time as this one. - Abby, Fall 2023

6. Deep Learning for Computer Vision, COMS 4995 W (See Shreyas's note above regarding the course number)

a. Professor: Peter Belhumeur. The course is an introductory course to Deep Learning with a focus on Computer Vision applications (syllabus and assignments link). There is no textbook, exams, recitations, or attendance required in the course. The class assumes no prior DL experience, just an understanding of basic ML concepts and Linear Algebra. Grades are based off of five assignments (60%, split evenly) and an open-ended final project (proposal: 5%, final writeup + video: 35%). Prof Belhumeur is funny and clearly knowledgeable of the topic, but the course wasn't heavily theoretical (i.e. less theoretical than Stanford's CS231N class). The first three assignments are Python/numpy implementations of algorithms (a solver for SVMs, multilayer perceptrons etc) and the last two are Keras/Tensorflow/PyTorch based. They don't give you a lot of time for the assignments (around a week and a half) so I felt pretty crunched finishing the work on time (you get one slip "day" where you can submit an assignment late two days), and the second and third assignments were a lot more work than the other three. Make sure to follow the assignment spec to the T--I lost 30% of an assignment grade because I added backprop to the main function rather than implementing it as a separate method. Overall, I'd

recommend the class if you want some exposure in DL. - Young Sang Choi, Fall 2019

7. Analysis of Longitudinal Data (BIST P8157)
   a. Professor: [Srikesh Arunajadai](). This may not be an accurate evaluation of the course since the Instructor had several personal issues that disrupted 4-5 continuous weeks of the course and instruction was in-person. This course is definitely a "classic" biostats course where the instructor would write proofs on the board for 3 hours (no PowerPoints). It requires at minimum, knowledge of regression and would not recommend this course to those with only a basic level of statistics as it builds on existing coursework; the exams and homeworks were 50/50 applied/theoretical using R. HWs were fair (sometimes ambiguous with wording but he's ok if you explain your thought process and show your work) and workload was average but overall the consensus was the exams were difficult especially the final. Generally, the midterm average is in the 50-60% (?). I did have some nice takeaways. I can't remember if there is a textbook, but it's really what you put into it and mainly self-directed. If you're interested in time series and longitudinal data then it's likely worth it; I can't stress enough that Arunajadai is more theoretical (95%) and less applied (5%) in the classroom. He was not accessible outside of class but that could just have been circumstance. -Amanda, Fall 2019

8. Deep Learning in Biomedical Imaging (BMEN E4460)
   a. Professor(s): [Dr. Andrew Laine]() (BME) and [Dr. Jia Guo]() (Neurobiology / CUIMC). This class is an introduction to deep learning in computer vision with a medical bent. Topics include basics of CNNs (convolutions, backprop, etc.), GANs, and Vision Transformers. The course assumes pretty minimal background, and most of the material is self-contained. Two quizzes (5%), two assignments with a written and programming bit (30%), 24-hour programming take home midterm (25%), and a final project (40%). All work is with Python / PyTorch, with most students electing to use Kaggle / Colab. Grading is quite generous (midterm mean: 103.75% median: 105%, final project report mean: 33.25 / 35, median: 34.13 / 35), and they give you so much starter code for all parts of the course (assignments, midterm, final project) that it's not a ton of work. The meat of the course is a final project where you'll be assigned to a faculty mentor and a graduate student, where data and starter code will be provided for you. In past sections some students got ISBI/EMBC-type four page conference papers ([1](), [2](), [3]()) and [journal publications](). Overall, although I enjoyed the course I would recommend you take a deep learning course in the CS department unless you are really keen on working with Dr. Laine or Dr. Guo or you want a more mellow introduction to DL in medical imaging. –Young Sang Choi, Spring 2023

9. Introduction to Biostatistical Methods (BIST P6104)
   a. Professor(s): Sarah Ayton (Biostatistics). This class serves as a comprehensive introductory course for applying statistical methods in public health and biomedical contexts. Topics include basic probability, random variables,

population sampling, statistical inference, and simple/multiple linear/logistic regression. The course assumes you have little statistics background and is completely asynchronous. There are five assignments (50% total with the lowest assignment grade dropped), four timed quizzes (20%), and a final exam (30%). The assignments are clear-cut extensions of topics touched in lecture videos and the practice problems for each weekly module with some straightforward R coding thrown in. When taking the quizzes, you are allowed to access the class notes and materials. The quizzes are all multiple-choice, and you get two attempts for each quiz. The final exam is the same open-note policy as the quizzes, but it is a mixture of matching, short-response, and multiple-choice questions you can finish over a 48-hour period of your choosing. Grade distributions were not provided, but the grading of assignments is pretty lenient. Sarah Ayton also does an excellent job connecting concepts to tangible examples. Due to this and prior examples of the course's low-stress nature, this is a chill Quantitative elective that will equip you with an excellent toolkit for future biomedical research. -- Adit Anand, Fall 2023🧬🔢

10. Introduction to Probability and Statistics (STATGU4001)
    a. Prof.: Isabella Sanders. This was overall a great class. It was quite fast paced and the time slot of 3.5 h on a Monday night was a little unfortunate. Calculus is a prerequisite which was emphasized several times and calculus related questions are therefore specifically not answered in class. The slides are not that great unfortunately (handwritten in class or screenshotted morsels) and the textbook on the other hand is overly advanced and the problems are too complicated for what you learn in class, which we were warned about. Therefore, there is kind of a lack of practice problems and solutions. But because the syllabus is very straightforward, those were easy to find elsewhere. There was also no practice midterm or final. But the questions on the final were quite similar to the questions in a game of Jeopardy which we played in the last class. Classes were not recorded. Because it might be helpful: during the final weeks the possibility of obtaining extra credit by submitting personal in-class notes was announced, which would have been good to know in the beginning to take notes digitally or in some other appropriate format. There were only 2 TAs for about 130 people and even if they did a fantastic job it was probably a lot for them. The final project was very extensive. Some parts of the video introduction to R did not work, it might be helpful to supplement. None of the examples are from the biomedical field. For people without a CS or maths background, like me, this could be quite the challenging class but the grading is not super strict, you will get help if you ask the TAs and it is a very rewarding experience because you really learn a lot and feel a little proud of yourself if you make it. Laura Fall 2023

11. Mathematics of Machine Learning and Signal Processing (COMS 4995)
    a. The class covers really all the math you need for basic ML and signal processing. I took the class for the signal processing part, but ended up learning quite a lot about ML. The lectures were a bit dry, but slides were quite helpful. He records

all his lectures so you can watch later. It has a midterm and final project, and 4 or 5 homework assignments. The homework are easy proofs from lectures in class. The midterm was to implement the S4 from scratch (previous years was the transformer). The final project is anything that uses information used in this class, he also gives you quite a lot of project ideas. The Professor, Dr. Beigi, is very nice too. Very accessible through email, and he's always happy to meet with you to talk about your project. I can't say it's easy A, but definitely not too challenging. — Vivian 2023 Fall

Qualitative (Qual)

1. Special Topics in Biomedical Informatics: Intelligent Decision Support: History, Paradigms, Applications, BINF G4008 001
   a. Professor: Lena Mamykina. Taking this my first semester definitely made me feel as though I was in a little over my head, but in actuality it was a great headfirst dive into paper reading and analysis that has shaped how I approach informatics papers and research in general. The readings were definitely bulky and LONG (I don't know how things have toned down this semester, but it was definitely a bunch), but ended up being at least ~manageable~ and gave me a lot of resources I've used for other papers since. Overall, Lena is really receptive and understanding. We even had a long extension on our final paper (either 1 or 2 weeks later). Getting to pair up and feel equally interested and lost with a bunch of other students was actually quite exciting because we got to just try and act as though we were experts on decision support and cover a lot of different aspects of the field. – Sal Volpe, Fall 2019
2. Special Topics in Biomedical Informatics: Interrogating Ethics and Justice in Digital Health, BINF G4008 002
   a. Professors: Noémie Elhadad and Sandra Soo-Jin Lee. I don't have insight into how this format will change or how it differed from the previous offering that Sal detailed below. This is a great light-workload class that's relevant to just about any area you might go into, so I'd recommend taking it in your first year if you can. The majority of the work is readings (probably 60 pages per week) and responses. I found the readings easy to get through and pretty thought-provoking compared to other ethics classes I've taken. We had two final projects: an op-ed along with workshops held by an organization called The OpEd Project which was… a rocky experience so I wouldn't be surprised if this was scrapped, and a more "traditional" final project that wasn't too big in scope. Two notes: 1) because it's a small class and there's so much discussion, the class can be pretty draining. I enjoy talking about these issues but if you don't feel comfortable doing that it can be difficult. The final project presentation kind of ended up turning into a 1-hour oral exam, which was challenging to get through. 2) the expectations around projects/reading responses are pretty vague. Definitely ask for feedback and guidance early and often. - Chris Allsman, Fall 2022

3. Special Topics in Biomedical Informatics: Ethics and Fairness in Digital Health, BINF G4008 002 *(no longer offered)*
    a. Professor: Noémie Elhadad. For its first semester being taught, we got about 30% done in-person before the pandemic sent us all home, but it definitely stayed at least a fairly engaging course. Readings and discussions really centered around ethical practices in ML and how they apply to various industries and applications like recommender systems or AirBnB AI. The readings definitely weren't too intense and frequently we would diverge into more philosophical and moral discussions than staying too grounded in the paper, but Noémie and Tony (the TA) worked hard to get us back on course or at least staying in the domain of ML in Healthcare. We got to spend a lot of time on our own working toward a DIY final project for a proposal, lit review, whatever-you-made-of-it along with a final presentation. The highlights were definitely the guest lecturing and in-person learning. I'm interested to see how Noémie will take this in the Spring if she does it again. – Sal Volpe, Spring 2020
4. User Interface Design, COMS W4170
    a. Professor: Brian A. Smith. This course's waitlist fills up very quickly, so definitely try to be proactive (email Professor Smith) and attend the initial lectures. The course content covers topics touched upon in Research Methods (e.g., user-centered design, qualitative research methods, etc.) and dives into how to apply these concepts in the context of software application development (more of a UI/UX emphasis). Professor Smith is an excellent lecturer who consistently provides clean explanations and is not afraid to revisit concepts. The class also teaches how to develop web applications using HTML, CSS, JavaScript (vanilla JS, React), and Flask. There is a semester group project (teams of 3-5) focused on identifying a campus-related or college student problem space to develop an application for via user-centered design. The project has 5 different checkpoints during which you get feedback from a designated TA mentor. Overall, the project is not too strenuous as long as you ask the TA mentor to clarify the expectations and provide feedback on potential deliverables. There are also 5 homework assignments focused on coding and user-centered design that are reasonably spaced out during the semester. Attendance does factor into the final grade. In my opinion, this class has engaging content taught by a committed professor who is oftentimes receptive to student feedback! Professor Smith also repeatedly noted that grading is pretty generous. - Adit Anand, Fall 2024 💻🤝

Information Technology (IT)

1. Programming and Problem Solving, COMS 4444 W
    a. Professor: Ken Ross. I took this class more as a way to meet the degree requirements. It's a fun class where you solve open ended problems in groups, and go through multiple iterations of brainstorming, coding, and writing up your solution. There aren't many classes like this so it can be a good class to take for

that reason. The class is taught in Java which may or may not be a good thing. For me it made it difficult to be a part of the assignments as I could help brainstorm, but couldn't really code up anything. I had taken a class in Java a few years back but because I haven't used it consistently I kinda fell off. The class grade is based off of participation, and also peer reviews, so it's not hard to get a good grade in the class. -Oliver Bear Don't Walk IV, Fall 2017

2. Introduction to Blockchain Technology, ELEN 6883
   a. Tbh, I took this class because none of the other "approved" DBMI IT courses were available, and I just needed to take something. It wasn't too bad difficulty wise. The homeworks were a bit challenging, but overall I learned a ton about a topic I knew absolutely nothing about prior. Brb, gonna go dump my whole stipend into Bitcoin. - Jonah Einson, Spring 2020

3. Data visualization: QMSS G4063
   a. I wanted to take this class, but it's open to QMSS students only with a waitlist for external students. I tried twice to get into this class, but was never able to (see above statement). Looked pretty good though.
   b. Professor: Thomas Brambor. This was a very easy class – once per week lectures and 4 relatively unchallenging homework assignments. The homework assignments and the final project (group project designing a website around the data analysis of your choice) are not difficult, but do require some creativity and careful thought of the design principles discussed in the class. I think you'll get out what you put in (so if you pay more attention, you'll learn more). The class is mostly taught in R, although you can do the homework assignments in any language (they are easiest in R). – Apara, Spring 2024

4. Exploratory Data Analysis and Visualization: STAT GR5702 (Prof Joyce Robbins)
   a. This is a required course for DSI Masters students, so it can be hard to get into, but the professor was really open to working with me to get in the class and to make the work I did valuable for my own projects/research. It was valued that I could bring in practical examples from an applied domain (healthcare). The course is more than half in R with ggplot2, and at the end goes into interactive web viz with D3. The course material/topics are very foundational and not difficult at all, I found it helpful but also sometimes a bit (too) simple. Although I did learn some important underlying principles that will inform my research/design. The workload is light, with ~5 partnered problem sets, MC midterm/final, and a big viz project that is very flexible. If you can get in, and need an IT class, and use viz in your work--this class sounds like a good fit! --Adrienne Pichon, Spring 2020

5. Databases: COMS W4111
   a. Prof: Alex Biliris. If you're looking for an IT course, this one is offered every semester and has a number of sections so it may be hit or miss (it can also be difficult to get in; just join the long ~200-person waitlist as early as possible). Biliris has a kooky gruff but somehow likeable delivery and the class isn't that bad -- just 4 HWs, midterm, final, and project with many small components spread over several deadlines. Despite what's on the course directory, you DO NOT

have to know Java, the project is Python/SQL-based. For the [2-person team-based] project you can opt for the applied track and build a database backed/focused interactive website (think login administration/authentication, multi-search page, and lots of SQL queries) or you can go for the more theory-based track (expanded data modeling with Entity-Relation diagrams, some of this is covered in Symbolic Methods). I chose the applied track and it wasn't bad except if you don't know Flask/Jinja beforehand, you'll have to learn that on the fly (not a big deal). In fact, aside from SQL, all other programming is self-taught. All of the material needed for HW and exams is covered in lectures; the reading is optional. The project and midterm were fine, the final exam was considerably tougher (mean grade was ~52, but the class is graded on a curve). Consensus among DBMI folk I've spoken to is that the first half of the class is fine (just keep chugging along), the second half is less interesting/intuitive (think: storage indexing/optimization or calculating the number of disc accesses to perform X search). - Harry, Spring 2021

6. Analytics on the Cloud, IEOR 4526
   a. Prof: Hardeep Johar. This is a fantastic class that covers the basics of functional programming in Scala (with a little bit of Python too), how to use Apache Spark and the Google Cloud Platform for parallel/cloud computing, and includes applied examples of things like NLP, machine learning, SQL database querying, streaming, and graphical models. Professor Hardeep is fantastic at explaining how everything works, this was a hybrid class format and he lectured/provided notes in Jupyter notebooks. The homework assignments could be challenging but were well structured (and I felt way more prepared for these assignments than I ever did with the Computational Methods or Acculturation to Programming labs), and had very interesting applied topics (tracking COVID cases in NYC, Citi Bike usage, Twitter sentiment analysis, etc.). If you have no previous experience with a strongly typed language there can be a steep and time consuming learning curve with Scala, and finishing the coding sections of the midterm and final exams in time was challenging, but the grading was very fair. I took this as a summer course so covering all of the content in only 6 weeks while also working full time was VERY intense, but I think it's also offered in the spring semesters which sounds much more manageable. You also have to email the IEOR department to prove you have equivalent pre-requisites (for me Acculturation to Programming and Statistics + Computational Methods + work experience as a Data Analyst was enough) before they will accept you into the class. -Kristen King, Summer 2021

7. Data Science I, BIST P8105
   a. Prof: Jeff Goldsmith. This class is a must-take if you think you need any semblance of training in R or with the tidyverse (also minimal training in Git). Logically structured, it walks you through all the need-to-know steps for data cleaning and processing, and basic statistical techniques. Labs are incredibly do-able (usually fast) and use real raw data from various web sources so you will

get your experience data cleaning; you are also required to set up a GitHub repo where you post your labs and submit the assignment link. The final project is a big part of the class, and YMMV depending on who your group consists of, but it's mostly MPH and some MD students in the course so not the worst. The final project is a soup-to-nuts example of tackling a research question: determine your data source (up to the group), define your question(s), do your analyses, and create a web page to display your project. Jeff is a really good prof (very engaging, very responsive to emails/questions etc.) This class is probably incredibly easy for most people with experience in R, but for those like me who never developed good coding practices before, it was a gamechanger. -Courtney, Fall 2023

b. Second Courtney's rec (we took it together). I don't have much use for R in my research, but it's a generally helpful class for good coding practices, github use, etc. Also a very practically oriented class, good for those who learn by examples and getting your feet wet. I've used some things I learned in the class for my research just because the data organization skills are soo pleasing. - Kayla Schiffer-Kane, Fall 2023

Other Courses

1. Great Presentations, COMS 6915 E
   a. Professor: Janet Kayfets. Took this class at the behest of my PI. Not a ton of work, but expect to give a short presentation (1-5 minutes) every class. It's kind of fun, but Janet can be absolutely ruthless when she critiques you. That being said, I learned a lot about how to not suck so badly at presentations. -Oliver Bear Don't Walk IV, Fall 2018
2. Foundations of Data Privacy, IEOR 8100 / COMS 6998
   a. Professor: Rachel Cummings. The class title aside, this is a class that's specifically about differential privacy (with some applications to distributed systems + federated learning). I only audited the course, but it's *incredibly* theory heavy (coming from someone who considers themselves as having a decent theoretical background). Rachel also teaches a different class on Policy for Privacy Technologies: I'd recommend that one if you're interested in data privacy more generally, as it's much more applications-focused. You probably don't want to take this class unless you're interested in differential privacy or federated learning research, but do look into Policy for Privacy Technologies (which does count as an IT objective, I believe). - Chris Allsman, Fall 2023

DBMI Courses

1. Computer Applications in Health Care & Biomedicine, BINF 4001 G
   a. Professor: Nicholas Tatonetti. I'm pretty sure if you're here you have to take this class, so I'll focus on how to handle it. It's not a ton of work, but you'll get more out of it if you take the readings and lectures seriously. I think you can get a good grade in this class by just skating by, but the real beauty of the course is that you get a good intro to all the research being done in the department. If you're looking for rotations then come no further. This class allows you to learn about all the different aspects of the field and figure out what you might be interested in.-Oliver Bear Don't Walk IV, Fall 2017
   b. Professor: Katerina Schultebraucks. Just providing a more contemporary review, but I echo most of the above sentiments. We were entirely on Zoom, which may or may not continue to be the case as the pandemic continues, so do what it takes to set yourself up for success (i.e. be free of distractions, etc.)- it's a survey of a bunch of different lecturers, so some may be more engaging than others. The midterm paper is incredibly doable, they really set you up for success (i.e. you submit a proposal, an outline, and the final assignment, so many opportunities to figure out what they're looking for). The open-note final is also super doable- but definitely try to pay attention to what is actually *said* during the lectures and not just what's written! -Courtney, Fall 2021
   c. Professor: Gamze Gursoy. This is a fairly low stakes, classic seminar course. Twice a week, some person from DBMI/Sys Bio will come to give a talk, and the course is well divided into different domains (like general, bio, clinical, consumer health etc.) with talks ranging from single cell omics and microbiome to phenotyping and ontologies. It is a solid start to an MA/PhD program because it shows you the breadth of what is out there and what you can do in the field. The final "project", which is just a proposal presentation for a hypothetical project, is also not a terrible amount of effort. Just show up, learn, and present at the end of the year. - Aziz Zafar, Fall 2023
2. Computational Methods, BINF 4002 G
   a. Professor: Adler Perotte. I think this class is kind of in flux right now. Adler took over this last year, and I was one of the TAs during the course. Adler really knows his shit. Like this guy is wicked smart, and knows the math of what he's talking about inside and out. The course is meant to be an overview course and covers a lot of the methods you might use in BMI research at a high level, with the possibility of going more in depth using the readings or Adler. Because I wrote some of the homeworks I can tell you that they were a little hard. They required a solid grasp of coding (all done in python and using pytorch) as well as how to follow the workflow of CS assignments. They're more applied than theoretical, but do require students to implement some algorithms from the readings or class. So the course can be a lot of work, especially with homeworks every week, a

midterm, and a final project. I'm sure the course will change a lot for the next iterations, but you'll still get a good overview of the algorithms and methods used in BMI, with the ability to learn more if you like. To do well in this class work in groups and start the homework early.-Oliver Bear Don't Walk IV, Spring 2019 (TA)

b. Professor: Adler Perotte. This class is HARD. In a good way-- you'll learn a LOT, but you're touching the surface on many topics ranging from linear algebra to neural nets to DAGs. Full disclosure, I hadn't taken a calculus or algebra class in almost 15 years prior to taking this course, so the learning curve was quite steep. There are readings each week that are more theory which you then apply during labs. We had to turn in our reading notes as part of our homework so it was more effort than I'm used to for a grad class, but it did force you to read (and reread) before the lectures. The labs can be difficult (usually based on Python and PyTorch) but you will definitely learn (and our TA, Michael Zietz again, was excellent). The final project was a bit intimidating for me, but Adler provides excellent feedback and advice if you reach out. I felt like this class gave me a good idea of how to understand more detailed research/concepts in the field, but I think I'll need to take this course a few times before I'll really be able to master some of the topics. - Lauren, Spring 2021

c. Professor: Amelia Averitt. This was the first time Amelia (DBMI alumna) taught this course; she totally revamped it, and the dedication to creating a cohesive, thorough, and interesting syllabus/lecture content totally shows. As one with almost no computer science experience, I was able to keep up in the class just fine; there is a logical flow of topics from the start to finish of the course, and each lecture is jam-packed, but in a good way. Reading responses (whether summary notes or just commentary on the readings) are required for each class, but I found it a good way to summarize and interact with the content prior to lectures which was very useful for some of the denser topics like neural networks. Labs are (almost) weekly, and made brand new for this year (thanks Undina!), so some elements are likely to change over future iterations based on feedback from this year, i.e. implementing using SciKit Learn packages versus implementing from scratch. There was a hand-written midterm, which was a bit lengthy and definitely requires preparation. The final project doesn't enter the picture until the second half of the semester, but it comes at you fast- so definitely recommend giving it some good thought even before your proposal is due. Overall, phenomenal course, challenging and requires a lot of time, but manageable. -Courtney, Spring 2022

d. NOTE: **Retitled to Machine Learning for Healthcare**, BINF4002G. Professor: Amelia Averitt. I think this course is good for people who have very little background in coding, ML, and linear algebra, because it is more of a "brush-stroke" ML course. The workload was very intense, since there were reading responses regularly (which is okay) and very (almost extremely) long and thorough labs. There is also an expectation that students *must* participate in

class, every single lecture. I recommend engaging with the material for that participation grade, since it is a lot of the grade. The course may change a LOT next year, since Dr. Joshi is teaching it. In any case, this is not what I would consider an easy course - be ready to spend a decent amount of time every week, and make sure to start thinking about your project well in advance. - Aziz Zafar, Spring 2024

3. Symbolic Methods, BINF 4003 G
    a. Professor: Chunhua Weng. I think the course has changed a lot since I've taken it. Not to throw anyone under the bus, but Jimmy Rogers would be a good guy to ask about it. However, I know this course is useful for going over the ontologies and some technologies that are useful for research in BMI. It's also not a ton of work, and the final project is pretty open ended. But I don't think you're allowed to use deep learning in the final project anymore. -Oliver Bear Don't Walk IV, Spring 2018
    b. Professor: Chunhua Weng. I took this course in my first semester as a DBMI student and it was HARD as someone who hadn't coded in a loooong time. That being said, knowing the ontologies and mapping tools available was incredibly useful for using/understanding things like OHDSI or anything that relies on clinical data in general (and I would argue this is quite useful for NLP as well). Even though I have a decent background in ICD9/ICD10/SNOMED hierarchies, this class made that knowledge applicable because I understood how these concepts were stored in clinical research databases. The labs were challenging, but our TA (shoutout to Michael Zietz!) was excellent and provided some great walk throughs. I also learned how to use SQL for the first time, which, while I'm still a novice, was very useful for programming in general. The final project for this course is a great way to get used to applying this kind of knowledge, and Chunhua gives fantastic and useful feedback on your proposals and your final submission. If I had taken this class after Computational Methods, I think I would have been able to do a lot more with the final project-- because I was "new" to coding, I could problem solve to a certain extent because of my clinical background, but couldn't really implement much. I think taking this class after Computational Methods will allow you to really flex your problem solving muscles a bit more, but I definitely learned a lot despite that. - Lauren, Fall 2020
    c. Professor: Chunhua Weng. This is a first semester class, mandatory pretty much for every MA and PhD student in DBMI. I think this is a good class, and Chunhua is a lovely person. However, this is a very "clinical" course – ontological mappings, data harmonization, and common data models are all quite relevant to the Clinical Informatics students, but not so much for Bioinformatics students. Still, if you have to take it, make the most out of it and learn whatever you can. Course content is fairly manageable, and the workload is not super intense with the exception of the course project (which I feel should require a decent amount of work). Chunhua does expect a decent project, so do spend a few hours each week on it – don't wait until the end. - Aziz Zafar, Fall 2023

4. Research Methods, BINF 6002 G
    a. Professor: Lena Mamykina. I have mixed feelings about this course. I learned a lot about how to do research in general, qualitative methods for doing research and implementing new models, and writing a grant proposal. I think these aspects are incredibly important for doing well in research, however I can't see myself using the qualitative methods much given the more computational nature of my work. I'll elaborate more on that later, as I don't think it's useful for everyone. Lena is a boss, she really knows how to teach this course and is good at conveying her knowledge. The course is a good amount of work with readings, papers, discussions, etc, but is more back loaded with the final paper and final. I think the readings are useful for making you question why and how you're doing research which was a big awakening for me (galaxy brain meme). Ok, now getting into things that might be more useful for computationally oriented students. A large chunk of the course is focused on qualitative methods and I do think it would be useful if I were going to take my research a step further and implement my models, but as the focus of the department stands I can't see this happening. The only reason this is a big deal is because the final paper (a grant proposal) needs to have been influenced by what you learned as far as qualitative methods go. The department feels like it's in a bit of a change moment right now with balancing a need for both qualitative research approaches, and more computationally focused approaches.-Oliver Bear Don't Walk IV, Spring 2019
    b. Professor: Lena Mamykina. So first thing's first, when Lena says to seriously work on the research proposal progressively throughout, please listen! This class was a tough one for me because as Oliver said, the focus seemed to be on qualitative methods and prospective studies. It was tough to understand exactly how to cast more upstream quantitative methods research into the mold of a more qualitatively focused R21 grant proposal. That said, you should definitely talk to the TA and Lena to work with them on how to reorient your project into an intervention if you are in the same situation. I learned a whole lot in this course about all the different study designs (prospective, retrospective, etc.). I'd say this is very useful stuff for understanding research throughout biomedical informatics and beyond. This class (as a more methods quantitative person) forces you to think about the mechanics of how your methods would be operationalized and all of the problems you would encounter (such as measuring the impact of your CDS intervention). In this way, it was incredibly valuable for me but also probably on the harder side in terms of reorienting my thinking. One thing to watch out for is the requirements of the R21. I was somewhat confused throughout the semester on what kind of intervention (and what counted as an intervention) we had to include in the grant proposal. Ask questions early and often (and talk to Lena and the TA to clarify these things)! Don't be like me and put this off until the last weeks - Shreyas, Spring 2019

c. Professor: Lena Mamykina. Ditto what Oliver and Shreyas said. This class is definitely a shift in thinking for people who aren't familiar with qualitative research. There is a lot of work involved: there is a significant amount of reading but the assignments allow you to get quite creative! I will say that taking it at the same time as Computational Methods was quite stressful for me as someone who was re-learning quant methods and learning qual methods on top of it. Start working on the project early and ask lots of questions. Our TA was fantastic (shoutout to Adrienne Pichon!) and really helped crystallize some of the lecture content into more practical applications. Lena knows so much about this topic-- she wants to make sure you understand what you're doing and why you're doing it. I found the grant writing assignment to be helpful (received some excellent feedback), but I had already taken two grant-related courses at Columbia so I had some prior knowledge (funding & grantsmanship with Jamie Rubin in the spring and a grant writing course with Diachi Shimbo during CSRI). - Lauren, Spring 2021