Towards an AI Course Based on Neural Networks

Michael Wollowski

Rose-Hulman Institute of Technology, Terre Haute, Indiana, USA
wollowski@rose-hulman.edu

Abstract

In a prior paper, we argued that Artificial Intelligence (AI) should be placed on a different foundation, one based on pattern recognition and feature learning rather than symbol manipulation and feature engineering. In this paper, we provide a proof of concept of an AI course that follows that proposed approach. Students study how these systems become so incredibly powerful through machine learning of features and through pattern matching. Students learn how those systems represent knowledge and they study their currently limited reasoning abilities. Students spend time discussing the accomplishments of current systems, positive as well as negative and they study the projected impact of anticipated systems. In this paper, we give a brief argument of why one would want to offer such a course. We present a detailed outline of the contents of such a course, together with learning materials and their proposed use. We summarize relevant anonymous student feedback and offer a subjective evaluation of the pilot course.

Course Materials —
https://www.rose-hulman.edu/class/cs/csse313/schedule/

Introduction

In the final report of the CS2023: ACM/IEEE-CS/AAAI Computer Science Curricula (Kumar et al. 2024), in particular, in the portion relating to AI, Neural Networks (NNs) are taking a backseat to symbol-based AI. In (Lawless et al. 2024), we propose that AI be placed on a foundation of Neural Networks, feature learning and pattern recognition, rather than on symbol manipulation, heuristic search and feature engineering. Looking back at some early work in this field, we see that some researchers were aware that feature learning by machines and pattern recognition, the sort of things that are the hallmark of modern NN-based systems, would be necessary to develop powerful AI systems. For example, Minsky stated in (Minsky 1968): "Today, machines solve problems mainly according to the principles we build into them. Before long, we may learn how to set them to work upon the very special problem of improving their own capacity to solve problems." Selfridge et al. (Selfridge and Neisser 1963) wrote "The most important learning process of all is still untouched. No current program can generate test features on its own. The effectiveness of all of them is forever restricted by the ingenuity or arbitrariness of their programmers." Dreyfus, in (Dreyfus 1965, 1972) argued for the importance of pattern recognition in order to develop AI that operated like humans.

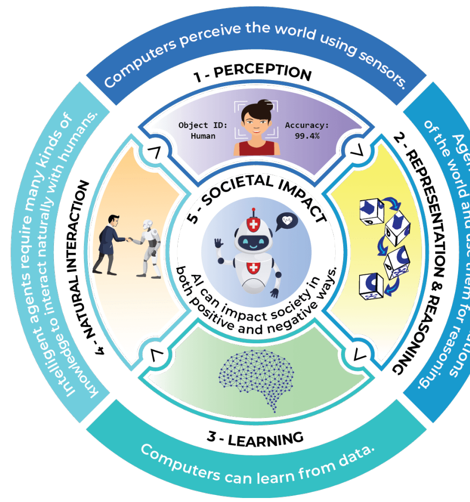

Figure 1: Five Big Ideas in Artificial Intelligence.

In the late 1960's, the field did not have sufficiently powerful computers and algorithms, nor the data storage capacity or digitized data to develop powerful NN-based systems. We believe that AI researchers pursued the next best thing: symbol-based AI systems. We believe it is time to change course and make NNs the basis of AI and develop a theory and learning materials around them.

In this paper, we provide a proof of concept of an AI course that follows our proposed approach. This course covers the big areas of AI, as documented by AI4K12 and displayed in figure 1.

The areas shown in the poster are independent of the approach taken towards AI. However, looking at the specifications of an AI course as found on page 84 of CS2023 (Kumar et al. 2024), we do cover some areas, in particular an introduction to AI, Machine Learning, knowledge representation and reasoning and society, ethics, and professionalism

(SEP).

- ”AI-Introduction (4 hours)
- AI-Search (9 hours)
- AI-KRR (4 hours)
- AI-ML (12 hours)
- AI-Probability (5 hours)
- AI-SEP (4 hours - integrated throughout the course).”

To understand the difference between symbol-based and NN-based systems, we do cover a very limited amount of search. We do not cover probability.

Motivations for Offering the Proposed Course

In this section, we provide some arguments about why one might base a general AI course on NN-based systems.

1. `Powerful Tools to Develop Powerful Systems.` We believe that one should introduce students to powerful tools, so they can accomplish amazing work. Currently, NN-based systems are more powerful than symbol-based systems.
2. `Motivation.` Students are fascinated by and curious about NN-based systems. Systems like the AlphaGo set of systems and ChatGPT made a big splash and students are curious to learn more about them. As such, motivating the course materials is fairly straight-forward.
3. `Needs.` As LeCun points out in (Stebbings 2001), neural networks are not nearly as powerful as they ought to be. As such, there is demand for research in NN. As LeCun also mentions in (Stebbings 2001), we will see applications based on NN. This requires engineering talent.
4. `Informing the Public Debate.` In 2023, several organizations expressed concerns about the state of the art and expected progress in AI, see (Future of Life Institute 2023; Center for AI Safety 2023; Bengio et al. 2024; AAAI 2023). We need domain experts who are capable of informing decision makers and the public about the power and limitations of modern AI systems and who can give sage advice about future directions of AI as well as how to steer it in a good direction.
5. `A Better Way to Make AI.` One might argue that all of the above are just fine and should be taught in a Machine or Deep Learning course. As we argued in the *Introduction*, we believe that feature learning by machines and pattern recognition are a more powerful way of developing AI systems than feature engineering by people and symbol-manipulation. To be clear, researchers in our field were not concerned about existential risks until NN-based AI became powerful. Yes, there were debates on the use of autonomous weapons, but they were few and far between.

Course Characteristics

To begin, we like to address how this course differs from a Deep Learning course, something that is offered at many institutions. We were not able to study courses at other institutions in detail, only the course offered at our institution. The major component of our Deep Learning course asks teams of students to propose a term project which they then develop with the use of existing packages. Instead, our AI course aims for students to understand the inner workings of NN. This is part of the overarching goal to explain the nature of AI and how modern AI is produced. In other words, this course is focused on students understanding what AI is and how it works at a fundamental level. This course is very much in flux and it is by no means perfect. However, we make significant revisions and improvements every time we teach it.

In this context, we believe that it is necessary to develop new learning materials, those that conceptualize AI in terms of pattern recognition, feature learning and many other concepts related to NN. We started this process, but we are clearly just at the beginning and hope to collaborate with others so as to place AI on a foundation appropriate for modern, NN-based systems.

Next, we provide learning outcomes and describe assessment items. Our school is on the quarter system and our quarters have ten weeks of instruction. This course meets four times a week for 50 minutes each. This course is placed in the Junior year so that students have taken several software development courses, including a data structures course. Additionally, our students are expected to have some maturity. Below is the current catalog course description.

> In this course, we will study modern AI systems, their current accomplishments, positive as well as negative, issues surrounding their training, and their inner workings. We will formalize those systems as pattern recognizers and distinguish them from classical, symbol-manipulating AI. We will study how these systems become so incredibly powerful through data-driven feature learning. We will look at how they represent knowledge and study their reasoning abilities. We will additionally spend some time discussing the projected impact of anticipated systems and study the building of beneficial AI systems.

The learning outcomes are as follows. Students who successfully complete the course will be able to:

1. Explain the architectures and training regimens of feed-forward and convolutional neural networks as well as key network architectures used for sequential natural language processing tasks, including transformers.
2. Explain how key hyper-parameters, data sets, data curation, and the data itself affect the quality and performance of neural network systems. Explain ethical issues surrounding copyright of and bias in the data.
3. Explain pattern recognition, feature learning, knowledge representation and reasoning in key neural network architectures.
4. Explain key reinforcement learning algorithms.
5. Explain current applications of artificial intelligence.
6. Explain the anticipated impact of neural network systems, ethical issues surrounding the anticipated impact

and how to mitigate risk and ensure beneficial neural network systems.

7. Implement a feed-forward and a convolution neural network.
8. Fine-tune a large-language model.

The assessment items are described below; the percentages indicate the weighted contributions to the course grade.

- Projects. [50%] Around six projects related to coding, training, experimenting with AI systems. Some are pair projects.
- Essays. [10%] About seven summaries of key papers in the field.
- Cutting-Edge work presentation and report. [10%] Presentation of an AI application that is at the forefront of the field. This is a pair assignment. You may choose the topic. One of the topics you may choose is an evaluation of an LLM in a specific domain.
- Impact of AI presentation and report. [10%] Presentation of an assigned article concerned with the impact of AI. This is a small team assignment.
- Quizzes. [20%] Three in-class quizzes, covering the big items: FFnets, CNNs and Transformers and related architectures.

Course Contents

We have taught the course during this past Spring term. Based on that experience, discussions with students and anonymous student feedback, we made some significant changes, as discussed in the *Evaluation* section. The details below describe the course as we are teaching it during the current Fall term.

We use the following color coding to indicate the type of learning materials. The materials are listed roughly in order of appearance in the course.

- *Instructor class presentation*
- Writing assignment
- In-class discussion
- An assignment related to the use or development of software
- Student class presentation
- What is Intelligence? On the first day of classes, we like to address the following question: "What is intelligence and where do we find it?" Students are asked to form small groups and discuss the following prompt: "What is the lowest life form you consider intelligent and why?" After students discuss this prompt, we ask each group to tell us their top answer. The answers are recorded on the whiteboard and a brief class discussion follows each submission. This assignment helps students identify critical requirements or at least features of intelligent behavior. To offer perspective on human intelligence, we also discuss a quote ascribed to park a ranger: "There is considerable overlap between the intelligence of the smartest bears and the dumbest tourists." (Sienkiewicz 2023).
- *Five Big Ideas in AI*. On the first day, we additionally present and discuss the poster entitled "Five Big Ideas in Artificial Intelligence", developed by AI4K12.org (AI4K12.org 2020). This poster focuses on ideas that are fairly independent of an underlying approach towards AI.
- *Constraint Satisfaction*. The first programming assignment asks students to develop code to solve Sudoku problems. To support the assignment and to help us understand the symbol-based and NN-based approach towards AI, we introduce constraint satisfaction, backtracking, and optimizations through forward chaining and arc consistency.
- Turing Test. Students are asked to read and summarize key portions of Turing's paper entitled "Computing Machinery and Intelligence." (Turing 1950) It is a classic paper that enables us to discuss how to best assess systems considered intelligent.
- Turing Test Discussion. After the reviews of the "Computing Machinery and Intelligence" paper are due, we engage in a class discussion of the power and limitations of the Turing Test. Among others, we discuss limitations of a conversational test and discuss additional properties that make us human.
- Powerful Human-made Tools. When discussing the challenges of assessing intelligence, students are asked to form groups and produce a list of powerful engineered tools. This is to provide perspective of human acceptance of engineered systems that perform better than humans. Tools that are mentioned include cars, bulldozers, airplanes and pocket calculators. We mention people's discomfort when it comes to powerful engineered tools that exhibit cognitive abilities, such as Large-Language Models (LLMs).
- Backtracking Sudoku Programming Assignment. Students are asked to implement a backtracking Sudoku solver. This is a warm-up assignment in which students, primarily non-majors, may relearn some of the intricacies of the Java programming language. We build up solving Sudoku problems as requiring intelligence. We point out that the student newspaper used to contain Sudoku problems, suggesting that they are sufficiently challenging to our student population. In addition to writing code to solve 9x9 Sudoku problems, students are asked to write an essay reflecting on whether they consider their software intelligent. We provide scaffolding for file I/O and provide several standard Sudoku problems, including one considered the hardest. As part of this assignment, we additionally provide 16x16 as well as 25x25 and 36x36 Sudoku problems. An extra credit assignment asks students to implement Don Knuth's dancing links data structure (Knuth 2000), in an attempt to efficiently solve larger Sudoku problems.
- Backtracking Search and Intelligence. After the deadline of the Sudoku assignment, we engage in a class discussion in which we address whether our students consider the backtracking Sudoku player intelligent. This ties into our day 3 discussion of the behavioral nature of the Turing test as well as our day 1 discussion of the nature of

intelligence. An unwritten goal of this class is to help students understand the nature of AI and what it can bring to the table. The discussion of the Sudoku code is a first significant step in honing students' understanding of AI.

- Deep Blue and Brute Force Search. Students are asked to read and summarize key portions of the paper about IBM's "Deep Blue" (Campbell, Hoane, and Hsu 2002). This paper is assigned soon after the reflection of students' backtracking Sudoku solver and serves to continue the discussion about brute-force problem solvers that additionally follow a classical symbol-manipulation and feature engineering approach.
- Deep Blue Discussion. After the reviews of the "Deep Blue" are due, we engage in a class discussion of the engineering accomplishments of the Deep Blue team as well as the long history of computer chess and the impact of Deep Blue. We watch a Youtube video about the historic match(Computer History Museum 2005) and discuss the nature of the brute-force approach of Deep Blue. We discuss to which degree we may wish to consider Deep Blue intelligent, continuing our discussing of whether Sudoku is considered intelligent. While this course is focused on NN-based AI, we believe it is important for our students to understand the exact nature of classical AI, in additional to understanding NN-based AI. This is so that students may appreciate the power of NN-based methods.
- *Monte-Carlo Tree Search*. We cover this algorithm to provide background for the reading assignment of the AlphaGoZero paper; see below.
- *Feed-forward Networks*. This set of presentations typically last for about five class sessions. It includes work sheets and an in-depth introduction the programming assignment as well as the testing data. During class, we introduce perceptrons and their learning algorithm as well as feed-forward networks and backpropagation. We explain the limitations of a perceptron and how multi-layer feed-forward networks overcome those limitations. We discuss simple pattern recognition such as in an AND perceptron. We also discuss how a FFnet with a single hidden layer processes the input pattern so that the activation at the hidden layer becomes linearly separable at the output node. In class, we ask students to trace the learning algorithms for a given perceptron and for a very simple FFnet, one that implements XOR. This helps students understand the learning algorithms as well as the test data we provide. We demonstrate the use of various training data in class and we demonstrate and explain the testing of the FFnet that is to recognize the MNIST data set; see the write-up of the assignment. We cover the importance of data curation, the importance of selecting a relevant training set, bias, and out-of-domain errors. In essence, we address how to properly train a neural network.
- FFnet programming assignment. In this assignment, students are asked to implement perceptron and feed-forward networks and to experiment with several hyper-parameters to develop an understanding of their effect on neural network learning and performance. The assignment culminates in training an FFnet on the MNIST data set (LeCun, Cortes, and Burges 2012). We provide students with the paper that explains how to read the data. There is a certain amount of curation that students may perform to increase learning efficiency. The purpose of this assignment is to help students develop an intuition of the necessary size of a NN as well as good values of various hyper-parameters for successful training of a NN and how they affect each other and learning. We explicitly ask students to experiment with hyper-parameter ranges that significantly delay learning. We discuss the algorithms in detail during class, and work out some of the intricacies through worksheets. Among the scaffolding we provide is test-code as well as a procedure that visualizes the MNIST input in 2D.
- *Ethics: Bias and Copyrights*. This is an important class session. To discuss bias, we present the interactive article "Humans are biased. Generative AI is even worse" (Nicoletti and Bass 2016). We discuss the recent lawsuit of the New York Times against OpenAI (Grynbaum and Mac 2023) and we discuss hallucinations and whether they should be called falsifications instead, as suggested by (Emsley 2023).
- Transition from Feature Engineering to Feature Learning. Students are asked to read the New York Times article entitled "The Great A.I. Awakening" (Lewis-Kraus 2016) and to submit a write-up containing the top-five things they learned from reading this paper. This article gives an excellent and well-written account of the development and break-through of feature learning with regard to Google's language translation system.
- Discussion of Google Translate's Switch from Feature Engineering to Feature Learning. After the deadline of the prior writing assignment, we engage in a class discussion of the top few items from each student's list so as to ensure students understand the power of the feature learning approach over the feature engineering approach. We now have made the transition from discussing symbol-based approaches of AI to NN-based approaches.
- *Early Realization of the Importance of Feature Learning*. In this presentation, we capture the sentiments of early AI researchers of the importance of learning features rather than engineering them. We will place this material in the context of the sort of problems they attempted to solve as well as the computing power and digitized data available at the time. Key sources are (Selfridge and Neisser 1963; Minsky 1968).
- *Reinforcement Learning* We spend about three days introducing Markov Decision Processes, value iteration, policy iteration and Q-Learning. We cover these materials because they are essential in understanding AlphaGoZero as well as some of the discussions surrounding the impact of AI, in particular the potential of rogue AIs.
- *Convolutional Neural Networks (CNNs)*. We spend five days covering CNNs. We explain their architecture and

demonstrate their processing of data. In this context, we discuss pattern recognition and, especially how at the later stages of processing, the patterns become unrecognizable to humans. We provide a brief history of CNNs and explain the details of the back-propagation algorithm. In the context of studying LeNet5 (LeCun et al. 1998), we revisit data curation.

- AlphaGoZero. Students are asked to read and summarize the "Mastering the game of Go without human knowledge" paper (Silver et al. 2017). This paper describes an early and impactful system that achieved remarkable public success solely by feature engineering.
- AlphaGoZero Discussion. After the prior writing assignment is due, we discuss the performance of that system. We focus on feature learning, how reinforcement learning was employed and the pattern matching performed by the system.
- CNN Forward Pass Programming Assignment. In this assignment, students are asked to implement the forward pass of LeNet5 (LeCun et al. 1998) for MNIST character recognition. This is a new assignment. We are adding it based on conversations with students during the Spring term, as well as feedback from anonymous student course evaluations. During this past summer, we worked with a student to develop a proof of concept of this assignment as well as ascertain which sort of scaffolding we wish to provide. The purpose of this assignment is for students to get an under-the-hood view of CNNs and pattern recognition. LeNet5 is a pioneering CNN and it has a fairly simple architecture. Students will need to make minor changes to their code to read in the data. For the classifier portion of the algorithm, they are asked to use their existing FFnet. The algorithm will be discussed in class and explored through worksheets. For this assignment too, we provide test cases to ensure the implementation is correct and we provide an adapted version of the visualization software that we made available for the FFnet implementation of MNIST character recognition. Among others, the visualization software enables students to observe the processing of the CNN.
- CNN Backpropagation Programming Assignment. Since implementing a CNN from scratch is challenging, we decided to split this assignment into two. The visualization mentioned above would be meaningless without the backpropagation portion, hence students will reap the reward of visualization only after completing both assignments. We discuss the learning algorithm in detail during class, complete with worksheets.
- CNN Sudoku Training Assignment. To compare and contrast feature learning with feature engineering, we ask our students to train a given CNN to self-learn and play 9x9 Sudoku. This assignment is adapted from (Mustafa 2020). As part of this assignment, students learn about installing and using Tensorflow and Keras. Some students know about those frameworks already, but for some students this is new.
- *Knowledge Representation in NNs*. In this presentation, we explain and dissect knowledge representation in embeddings. We show how embeddings represent ontologies and present some work that shows how common sense emerges from embeddings. We discuss common sense and how difficult the engineering of common sense is.
- *Reasoning in LLMs*. In this presentation, we show and discuss LLMs reasoning ability as detailed in several resources (Zhang and et al. 2022; Hagendorff, Fabi, and Kosinski 2023; Russell 2019; Chen et al. 2023). Among others, we discuss its ability to solve logic puzzles, common sense problems, and reasoning problems requiring common sense. We highlight "chain-of-reasoning" explanations given by ChatGPT and other LLMs. This is an area of much interest and some progress. Once available, we will study OpenAIs Strawberry model (OpenAI 2024) to assess its reasoning and problem solving capabilities.
- *Future Network Architectures*. In this presentation, we discuss the limitations of LLM when it comes to reasoning. We discuss criteria and requirements for systems that reason, as posed by (Lenat and Marcus 2023; Lake et al. 2016). We discuss approaches that favor the combination of NN-based and symbol-based approaches (Center for Integrated Cognition 2023). We believe that advanced reasoning has to grow organically from within NN systems. To this end, we present work aimed at developing new network architectures, work that is actively pursued by (Stebbings 2001; Heikkilä 2023; Chen et al. 2023; Institute for Foundation of Machine Learning 2023). For example, in (Heikkilä 2023), LeCun is quoted as follows: "One of the main challenges to improve LLM reliability is to replace Auto-Regressive token prediction with planning. Pretty much every top lab (FAIR, DeepMind, OpenAI etc) is working on that and some have already published ideas and results. [Note: I've been advocating for deep learning architecture capable of planning since 2016]."
- LLMs. Students are asked to read and summarize the article entitled "What Kind of Mind Does ChatGPT Have?" (Newport 2023). This is a very accessible article that explains how LLMs work.
- LLM Discussion. After the prior writing assignment is due, we discuss key points of the article as well as the challenge of predicting the next word. We point out that characterizing LLMs simply as next word predictors does not do them justice. Among others, we challenge our students to predict the next word in examples taken from literature and technical writings. We compare student responses with those given by ChatGPT.
- *Hopfield networks*. We spend half a session on Hopfield networks, primarily to show students that there are networks architectures other than the two pass architectures in prominent current use.
- *NLP*. During this class session, we cover early work on natural language processing (NLP). We briefly summarize very early work from the 1950's (UNESCO 1959; The University of Washington 1958) and we give a brief demonstration of Eliza (Weizenbaum 1966) and explain

its architecture. We place NLP in the context of speech acts and discuss some of the challenges of NLP, surrounding various forms of ambiguity of natural language processing.

- *Networks Designed for Natural Language Processing*. During this set of five class sessions, we examine network architectures designed for natural language processing. The class materials lean heavily on the textbook "Speech and Language Processing" by (Jurafsky and Martin 2024). We cover Recurrent Neural Networks (RNNs), networks based on Long-Short Term Memory units (LSTMs), and Transformer architectures; culminating in discussing the "Attention" paper, see below.
- Attention is All You Need. Students are asked to read and summarize the paper "Attention is all you need" (Vaswani et al. 2017). This is a very technical paper; however, we do explain transformer architectures in class beforehand.
- Attention Paper Discussion. After the prior writing assignment is due, we discuss key points of the paper. Due to the highly technical nature of this paper, part of the discussion serves as a review of the materials covered in class. However, we additionally study the kind of pattern recognition performed by the system.
- Fine-tuning LLaMA Assignment. In this assignment, pairs of students are asked to fine-tune a Large Language Model (LLM) using an existing data set. Students are asked to fine-tune LLaMA 2 (Kaparthy 2023) to build a model of the TinyStories dataset (Eldan and Li 2023). This assignment is adapted from (Sridykhan 2023). We expect to further develop this assignment during the next summer.
- Cutting-Edge Work. In order to get a sense of the breadth of AI, as well as of current work in this field, pairs of students are invited to choose a research article, a research project, or an interesting AI application, study it and present the chosen project to class. Students need to seek instructor approval, primarily to ensure the scope of the selected project is well-chosen. As an alternative to this project, pairs of students are invited to conduct experiments with which to establish the power and limitations of current LLMs. Students are encouraged to select a fairly narrow domain and conduct well-chosen experiments aimed at ascertaining the usefulness and quality of the knowledge exhibited by their chosen LLM. Here too, students need to seek instructor approval to ensure an appropriately chosen domain. Among others, the cutting-edge work presentations enable students to perform an initial study of a topic which they may wish to pursue further, among others, through pursuing a senior thesis.
- Cutting-Edge Work Presentation. The cutting-edge work presentations are scheduled through-out the second-half of the term, typically on Fridays to end the week on a fun and interesting note. Pairs of students are asked to give a 12-minute class presentation, followed by 4 minutes of discussion.
- Impact of AI. Teams of 5 to 6 students read and discuss one of the following papers. They are asked to develop a slide-presentation that captures the key points of the assigned paper and to lead a class discussion of those key points.
  1. Bengio, Hinton et al. "Managing AI Risks in an Era of Rapid Progress" (Bengio et al. 2024)
  2. Henry Kissinger: "How the Enlightenment Ends" (Kissinger 2018).
  3. E. M. Forster: "The Machine Stops" (Forster 1909)
  4. Stuart Russell: "Provably Beneficial Artificial Intelligence" (Russell 2017)
- Impact of AI Presentation. The teams of students from the prior assignment present their slides and lead a class discussion on the presentation topic. To ensure an in-depth discussion, we ask students to spend about 10 minutes presenting key points and lead a 15-minute discussion.
- Alchemy and AI. For this assignment, students are asked to summarize the first portion of the paper that questions the early symbol-manipulation approach towards AI, entitled "Alchemy and Artificial Intelligence" (Dreyfus 1965). This paper focuses on pattern recognition and other cognitive processes pertinent to human information processing.
- Alchemy and AI Discussion. After the deadline of the prior writing assignment has passed, we engage in a class discussion of the paper.
- *HCI*. We cover this area in the context of natural interaction. This is an area with tremendous job-growth, see (Ellingrud et al. 2023) and a lot of engineering. We believe that our students might be interested in the engineering aspect of this field.
- *Internet of Things (IoT)*. We cover the IoT, to raise students' awareness of the untapped potential of AI. By generating data and combining data from many devices, many aspects of our lives can be automated by the use of AI.

Evaluation

So far, the response to the refocused course has been overwhelmingly positive.

Feedback from Colleagues

When we developed this course, we obtained positive feedback from two alumni, one is a PhD working in NLP and the other is a trailblazing PhD student in NLP focusing on developing and using foundational models. We additionally obtained positive feedback from two colleagues at other academic institutions, one is a logician who recently bridged the gap to NN-based systems the other one is an expert in human-machine teaming. Two colleagues in my department, both with expertise in AI, are tasked with reviewing course assessment documents for our AI course. Both gave positive feedback and during a department meeting at which we discussed this new course, we obtained positive feedback.

Feedback from Students

Perhaps most importantly, our Spring term students were very supportive of the new version of this course as judged by the numeric and written course evaluations. Even during the course, we received excellent feedback for improving this course. Among others, students volunteered to complete additional or more in-depth versions of the assignments. There was a lot of excitement, as judged by the number and kind of feedback we received from our students.

Next, we will summarize confidential student feedback as well as corresponding changes we made to this course.

Learning. Students were excited about their learning. Students explicitly mentioned the refocused learning on NN, the software development assignments and the readings as conducive to their learning. There were a few scattered remarks about adding more scaffolding to the software development assignments, by providing more detail in class, and by providing what might be called ”warm-up” assignments. Additionally, it appeared that those students who already took the Deep Learning class wished for more technical depth of the assignments. For various reasons, the software development assignments were not as developed as we would have liked them to be. However, we obtained a summer grant from our department that enabled us to revise the FFnet and CNN assignments so as to provide technical depth. Students are now asked to implement those two architectures in Java.

Strengths. Students indicated that the software development assignments as well as the readings are a strength of this course. Several students stated that refocusing the course on NN was a strength. Students felt that the course materials reinforced each other. Some students called out the class discussions as a strength of this course. Some students named the diversity of the materials, i.e. readings, discussions, exploration of topics of their choosing, as a strength of this course and liked the ability to perform some research and to become independent learners.

Improvements. Our students suggested to improve the write-ups of the assignments, refocus them on learning key items and expand them. As indicated above, this is exactly what we did. We were thinking of removing the ChatGPT assignment, however, a fair number of students stated that they liked it. One student had an excellent suggestion: to offer the investigation of the power and limits of ChatGPT as part of the cutting-edge presentations. This is exactly what we are doing now.

Changes Made Based on Feedback

Based on our own thinking as well as student feedback, we made the following changes for the current version of this course. Please notice that this paper does describe the revised version of the course.

1. Ask students to implement an FFnet for MNIST character recognition.
2. Ask students to implement a simple CNN, again, for MNIST character recognition.
3. Provide additional learning materials for all three major network architectures covered in this course: FFnets, CNNs, and Transformers. This includes lecture materials, in-class worksheets and where applicable, a more in-depth discussion of the challenges and requirements of the programming assignments. The latter includes an explanation of the test-data we developed.
4. Discuss data curation in more detail.
5. We removed a learning outcome and assignment related to assessing the power and limitations of LLMs. Instead we made that assignment an option for the ”cutting-edge work” assignment.
6. We added three quizzes, related to the three major network architectures. Since we have, and continue to provide, more and more technical details of those architectures, we believe that the quizzes help in learning about their details.
7. One change we did not make is related to the LLMs assignment. We are keeping it for now, hoping that we will have time next summer to modify it.

Conclusions

In this paper, we provide a proof of concept AI course that is focused on feature learning and pattern recognition. We provide a brief argument of why one would want to refocus an introductory AI course in this manner rather than simply ask for a deep learning course. We provide a catalog course description, learning outcomes and a description of assessment tools. The majority of the paper is dedicated to justifying and describing the various learning components of this course. Among others, we highlight where feature learning and pattern matching is covered in the learning materials. We end the paper with a subjective evaluation as well as by summarizing relevant anonymous student course evaluations. We welcome collaborating with colleagues who are interested and willing to refine the proposed approach and to develop and refine learning materials towards an AI course based on NNs.

References

AAAI. 2023. Working together on our future with AI. https://aaai.org/working-together-on-our-future-with-ai/. Accessed: 2023-07-20.

AI4K12.org. 2020. Five big ideas in Artificial Intelligence. https://ai4k12.org/resources/big-ideas-poster/. Accessed: 2023-07-20.

Bengio, Y.; Hinton, G.; Yao, A.; Song, D.; Abbeel, P.; Darrell, T.; Harari, Y. N.; Zhang, Y.-Q.; Xue, L.; Shalev-Shwartz, S.; Hadfield, G.; Clune, J.; Maharaj, T.; Hutter, F.; Baydin, A.; McIlraith, S.; Gao, Q.; Acharya, A.; Krueger, D.; Dragan, A.; Torr, P.; Russell, S.; Kahneman, D.; Brauner, J.; and Mindermann, S. 2024. Managing extreme AI risks amid rapid progress. *Science*, 384(6698): 842–845.

Campbell, M.; Hoane, A.; and Hsu, F.-H. 2002. Deep Blue. *Artificial Intelligence*, 134(1): 57–83.

Center for AI Safety. 2023. Statement on AI Risk. https://www.safe.ai/statement-on-ai-risk. Accessed: 2023-07-20.

Center for Integrated Cognition. 2023. AAAI Fall Symposium on Integration of Cognitive Architectures and Generative Models. https://integratedcognition.ai/fss-2023/. Accessed: 2024-01-18.

Chen, Z.; Weiss, G.; Mitchell, E.; Celikyilmaz, A.; and Bosselut, A. 2023. Reckoning: Reasoning through Dynamic Knowledge Encoding. In *Proceeding of the Thirty-Seventh Conference on Neural Information Processing Systems (NeurIPS 2023)*. New Orleans, LA.

Computer History Museum. 2005. 1997 Match of Kasparov against Deep Blue. https://www.youtube.com/watch?v=NJarxpYyoFI. Accessed: 2024-09-14.

Dreyfus, H. L. 1965. Alchemy and Artificial Intelligence. RAND Corporation Technical Report P-3244, RAND Corporation, Santa Monica, CA. Accessed: 2023-07-20.

Dreyfus, H. L. 1972. *What Computers Can't Do*. New York, NY: Harper and Row Publishers, 1st. edition.

Eldan, R.; and Li, Y. 2023. TinyStories: How Small Can Language Models Be and Still Speak Coherent English? https://arxiv.org/abs/2305.07759. Accessed: 2024-01-18.

Ellingrud, K.; Sanghvi, S.; Dandona, G. S.; Madgavkar, A.; Chui, M.; White, O.; and Hasebe, P. 2023. Generative AI and the future of work in America. https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-of-work-in-america. Accessed: 2024-09-14.

Emsley, R. 2023. ChatGPT: these are not hallucinations, - they're fabrications and falsifications. *Schizophrenia*, 9(52).

Forster, E. M. 1909. The Machine Stops. *The Oxford and Cambridge Review*.

Future of Life Institute. 2023. Pause Giant AI Experiments: An Open Letter. https://futureoflife.org/open-letter/pause-giant-ai-experiments/. Accessed: 2023-07-20.

Grynbaum, M. M.; and Mac, R. 2023. The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work. https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html. Accessed: 2024-01-18.

Hagendorff, T.; Fabi, S.; and Kosinski, M. 2023. Human-like intuitive behavior and reasoning biases emerged in large language models but disappeared in ChatGPT. *Nature Computational Science*, 4(10): 833–838.

Heikkilä, M. 2023. Unpacking the hype a round OpenAI's rumored new Q* model. https://www.technologyreview.com/2023/11/27/1083886/unpacking-the-hype-around-openais-rumored-new-q-model/. Accessed: 2024-01-18.

Institute for Foundation of Machine Learning. 2023. We are the National AI Institute for Foundations of Machine Learning (IFML). https://www.ifml.institute/national-ai-institute-foundational-research. Accessed: 2023-10-25.

Jurafsky, D.; and Martin, J. H. 2024. Speech and Language Processing (3rd ed. draft). https://web.stanford.edu/~jurafsky/slp3/. Accessed: 2024-08-20.

Kaparthy, A. 2023. Llama2 code repository. https://github.com/karpathy/llama2.c. Accessed: 2024-01-18.

Kissinger, H. 2018. How the Enlightenment Ends. *The Atlantic*.

Knuth, D. E. 2000. Dancing links. https://arxiv.org/abs/cs/0011047. Accessed: 2023-07-20.

Kumar, A. N.; Raj, R. K.; Aly, S. G.; Anderson, M. D.; Becker, B. A.; Blumenthal, R. L.; Eaton, E.; Epstein, S. L.; Goldweber, M.; Jalote, P.; Lea, D.; Oudshoorn, M.; Pias, M.; Reiser, S.; Servin, C.; Simha, R.; Winters, T.; and Xiang, Q. 2024. *Computer Science Curricula 2023*. New York, NY, USA: Association for Computing Machinery. ISBN 9798400710339.

Lake, B. M.; Ullman, T. D.; Tenenbaum, J. B.; and Gershman, S. J. 2016. Building Machines That Learn and Think Like People. https://arxiv.org/abs/1604.00289. Accessed: 2024-09-11.

Lawless, W. F.; Mittu, R.; Sofge, D.; and Fouad, H., eds. 2024. *Interdependent Human-Machine Teams*, chapter Towards a New Foundation for AI. Academic Press.

LeCun, Y.; Bottou, L.; Bengio, Y.; and Haffner, P. 1998. Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*.

LeCun, Y.; Cortes, C.; and Burges, C. J. 2012. The MNIST Database of Handwritten Digits. https://yann.lecun.com/exdb/mnist/. Accessed: 2024-09-11.

Lenat, D.; and Marcus, G. 2023. Getting from Generative AI to Trustworthy AI: What LLMs might learn from Cyc. https://arxiv.org/abs/2308.04445. Accessed: 2024-01-18.

Lewis-Kraus, G. 2016. The Great A.I. Awakening. https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html. Accessed: 2023-07-20.

Minsky, M. 1968. Machines are More Than They Seem. *Science Journal*, 4(10): 3–4.

Mustafa. 2020. Solving Sudoku using CNN. https://www.kaggle.com/code/mustisid/solving-sudoku-using-cnn/notebook. Accessed: 2024-01-18.

Newport, C. 2023. What Kind of Mind Does ChatGPT Have? *The New Yorker*.

Nicoletti, L.; and Bass, D. 2016. Humans are biased. Generative AI is even worse. https://www.bloomberg.com/graphics/2023-generative-ai-bias/. Accessed: 2024-01-18.

OpenAI. 2024. Learning to Reason with LLMs. https://openai.com/index/learning-to-reason-with-llms/. Accessed: 2024-09-14.

Russell, S. 2017. Provably Beneficial Artificial Intelligence. https://people.eecs.berkeley.edu/~russell/papers/russell-bbvabook17-pbai.pdf. Accessed: 2024-01-18.

Russell, S. 2019. *Human Compatible*. New York, NY: Viking Press, 1st. edition.

Selfridge, O. G.; and Neisser, U. 1963. Pattern Recognition by Machine. In Feigenbaum, E.; and Feldman, J., eds., *Computers and Thought*, 237–250. New York, NY: McGraw-Hill, 1st. edition.

Sienkiewicz, R. 2023. There is considerable overlap between the intelligence of the smartest bears and the dumbest tourists. https://www.velvetshark.com/til/til-smartest-bears-dumbest-tourists-overlap. Accessed: 2023-07-20.

Silver, D.; Schrittwieser, J.; Simonyan, K.; Antonoglou, I.; Huang, A.; Guez, A.; Hubert, T.; Baker, L.; Lai, M.; Bolton, A.; Chen, Y.; Lillicrap, T.; Hui, F.; Sifre, L.; van den Driessche, G.; Graepel, T.; and Hassabis, D. 2017. Mastering the game of Go without human knowledge. *Nature*, 550(19): 354–359.

Sridykhan, C. 2023. Train from scratch and Finetune an Instruct Llama2 model in PyTorch. https://medium.com/@cindysridykhan/how-to-train-and-fine-tune-an-instruct-llama2-model-in-pytorch-6cbe11de2b34. Accessed: 2024-01-18.

Stebbings, H. 2001. Harry Stebbings interview of Yann LeCun. https://www.youtube.com/watch?v=OgWaowYiBPM&t=27s. Accessed: 2023-07-20.

The University of Washington. 1958. Linguistic and Engineering Studies in The Automatic Translation of Scientific Russian Into English. Technical Report RADC-TN-58-321, The University of Washington, Seattle, Washington.

Turing, A. M. 1950. Computing Machinery and Intelligence. *Mind*, LIX(236): 433–460.

UNESCO. 1959. *Proceedings of the International Conference on Information Processing*. Paris, France: Unesco, 1st. edition.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; and Polosukhin, I. 2017. Attention Is All You Need. In *Proceeding of the Advances in Neural Information Processing Systems 30 (NIPS 2017)*.

Weizenbaum, J. 1966. ELIZA - a computer program for the study of natural language communication between man and machine. *Commun. ACM*, 9(1): 36?45.

Zhang, Y.; and et al., A. B. 2022. Unveiling Transformers with LEGO: a synthetic reasoning task. https://doi.org/10.48550/arXiv.2206.04301. Accessed: 2024-01-18.