**CIS 600: Research Methods**

**Individual Assignment**

**Fall 2022**

**Due Date: Sunday, November 13, 11:59 PM**

**200 Points**

1. **Attend at least two online workshops organized by CESR. See the information in the CESR Workshops document. Discuss five takeaways from the sessions you attended. Provide the following information:**
2. **Title of the session(s) you attended.**

Manuscript Preparation

IRB Application

1. **Discuss each of your takeaways.**

**Manuscript Preparation**

This workshop discussed the basic steps of writing a manuscript. From the very first step of defining your audience and selecting an outline to the end of the publication process, writing a manuscript is an especially important cycle in journal articles and other research publications. The main takeaways from this workshop are how to write a manuscript, the process, the steps when writing a manuscript, there are a total of after submitting a manuscript to a journal, the importance of peer-review, and how to manage the decisions made by the publication. There are ten steps in writing a manuscript before submitting it to a journal. The ten steps in writing a manuscript are as follows

**Step 1: Defining your audience and selecting an outlet**

When writing a manuscript for a particular journal, we should ensure that the journal readers will be interested in our topic and in the journal specialized on a specific topic/subject or whether it is a comprehensive journal. Based on the journal type, we will be able to choose if we will be writing a narrow manuscript topic or a broad topic. One should also consider the reputation and quality of the journal because that will determine your acceptance rate and the quality of your writing. Another consideration would be the timeliness consideration of the journal. Since it is ethically prohibited to submit a manuscript to other journals while being under consideration in another journal, knowing how long it takes a journal to get back to you about being accepted or rejected is especially important. It is also important to follow the guidelines provided by the journals we are writing for including the paper format, citations, and length of the manuscript.

**Step 2: Write the Methods**

The workshop explained the best methodologies are the two “cs” methods which are clean and clear. The clean method means that the study is based on appropriate, valid, and unflawed methods of sampling and the use of instruments, procedures, and analysis. The clear method means describing what you did in the study in sufficient detail so experts in the field could replicate the study to check the reproducibility and generalizability of the study. The methodology should also include describing the samples used in the study (demographic characteristics, and others), measurements of the DV, and research design and plan for analysis.

**Step 3: Describe your results**

When discussing your result, one should mention all the relevant results regardless of whether they support the initial hypothesis (statistically significant and nonsignificant items, small effect sizes, and so on). Do not avoid unexpected results that make you uncomfortable since they are important to the study. Later in the manuscript, the implications of the results will be discussed in detail.

**Step 4: Discuss your findings**

This section will be where we will evaluate and interpret the findings of our study. Findings will include statements that support our initial hypothesis if our initial hypothesis was correct or statements of nonsupport if our initial hypothesis turned out to be proven wrong. Discuss the internal validity and external validity of the sources we use in the study. Common internal validity includes imprecision of measures, the overall number of tests or overlap among tests, effect sizes, and other weaknesses of the study. We will also be discussing the limitation and the importance of our findings. If possible, we should provide a link to future research.

**Step 5: Write the introduction:**

In the workshop, they discussed the four foundations of introduction. These include establishing the current knowledge we have about the field, summarizing previous research, providing wider context, going over the background of the study and its importance, setting the stages for the present research, explaining the gaps in the knowledge, and presenting research questions, and lastly introducing present research, the purpose of the study, and the outline of the design.

**Step 6: References**

This section will demonstrate how your findings are related to earlier reports on the same topic. One thing we should also consider is assuming that the reviewers will be the authors of the papers you use as your reference. Make sure the format you use for your reference is the same as what the journal requires.

**Step 7: Write the abstract**

The abstract is written after you have written the rest of the manuscript and should consult the guidelines provided by the journal to have proper formatting and structure while staying within the word limit provided. The abstract will summarize the hypotheses, designs, and finding of the study. It will also represent the article in indexing databases.

**Step 8: Create the title page**

Make sure the title captures the reader’s attention and informs the reader of the content of the study. The title page should also contain all the information required by the journal.

**Step 9: Edit and review your manuscript**

Fix any mistakes made when writing the manuscript initially including grammatical errors, fixing the continuity of the manuscript, and other mistakes so it satisfies what the journal expects. Be sure to have a manuscript that is clear to understand and precise.

**Step 10: Circulate your manuscript**

Before actually submitting your manuscript to a journal, having a second eye to read your manuscript as a test drive will benefit you and help with some changes that should be made to the manuscript. After the final review based on the feedback different readers have given you, your manuscript can be submitted to a journal.

After the submission of your manuscript to a journal you initially had in mind, there are multiple processes before reaching the decision made by the publication. The first step after a journal gets your submission is the editor giving the manuscript a preliminary read which means checking for the proper formatting style, the type of article that is appropriate for the article, and if the content is within what the journal puts out. If the manuscript does not fulfill one of these requirements, the editor will inform the editor for the author to fix the mistakes. Otherwise, the manuscript will be accepted and go through the peer review process which has a different timeframe depending on the journal but will be anywhere between 2-3 months.

The peer review process consists of 2-3 reviewers which are selected based on expertise, attention to the perspective, and familiarity with the topic and the controversy in that topic. The reviewers will offer constructive criticism towards the author of the manuscript which will help the author better their writing and adjust to get chosen for publication. Peer reviewers will do what is known as a quick read which is scanning the manuscript from beginning to end to look for obvious flaws. This quick read will include reading the abstract, examining the full manuscript, scanning the paper’s heading, references, tables, and figures, and the quick read is finished by reading a page or two from each section of the manuscript.

After the peer review process, comes the publication decision which will determine whether your manuscript is rejected, invited to revise, and resubmit, or if it is accepted. Each decision has its reason for that decision. If a manuscript is rejected, the main reason for the rejection is that the manuscript is outside the area of coverage of the journal, contains flaws in design, methodology, analysis, or interpretation, and the manuscript is believed to make only a limited novel contribution to the field. If a manuscript is invited to revise and resubmit, it is not a guarantee it will be accepted but the author is advised to include more details, get another set of reviews, and make changes based on those reviews. A manuscript is very rarely accepted by a publication on the first try. It is usually accepted after the invitation to revise or resubmit a previously rejected manuscript has been revised. After acceptance, the manuscript will enter the publication process.

**IRB Application**

IRB is used to protect the welfare of participants of research under federal law and protect, in our case, the university from lawsuits from poorly conceptualized research or unethical activities. The guidelines of IRB application are they cannot review a submission until the PI (principal investigator) electronically signs it. PI cannot be students or adjunct faculty. Only the faculty mentor can be your PI. The time for reviewing an IRB application depends on the level of risk the research has, meaning the higher the risk/level, the longer it will take for the IRB board to review your IRB application. The IRB application will be done on the IRBNet website which will provide you with an application form you will fill out and use in your application process.

Before applying for an IRB, make sure you have the appropriate training required for your research. Make sure to have the necessary consent forms depending on the type of research method you are using to conduct research. Your PI should be mentioned in your IRB project and ask how they want to be contacted for future projects. On the IRBNet website, we can share full access to our PI so they can go through the application and sign appropriately. The researcher/student will have to sign the application as well. The person should expect IRB to take around 10 days to get back to you with their decision. Read all the reviews the IRB gives you, and all the edits made based on that review should be capitalized in all the documents and repeat the process of application.

1. **How do you plan to use the takeaways in your research in the future?**

The how-to-write-a-manuscript workshop will help me in future research projects because it explains how to draft manuscripts/research correctly and the steps involved in having a successful manuscript from beginning to end. Knowing each step will help in not leaving out key details that will set your manuscript above others so when it comes to the publication process, it will be invited to revise and resubmit and then get accepted.

If you are conducting research that will require human participants, make sure to go through the correct IRB training, choose a faculty member/mentor to be your PI, and fill out the IRB application form to not encounter any lawsuits and be unethical in your research.

*(40 points)*

1. **Create a video presentation on a topic of your interest. See the “Research Video Presentation” document for details. Submit the video file (.mp4). The video should be 8-10 minutes in length and must provide salient points about your topic. *(40 points)***

File: CIS600\_Individual\_Assignment\_Shiferaw.mp4

1. **State the title of your research project within 8-10 words. Use the best practices for creating effective titles discussed in the class. *(10 points)***

Activation functions: the mathematical aspects of creating new activation functions

1. **Identify at least ten scholarly articles related to your research topic. Use the library databases and Google Scholar to find the articles. Provide the citations in the ACM or IEEE format. Explain your rationale for selecting the articles. How did you discern whether an article you chose had high impact and was relevant for your research topic? *(25 points)***

[1] M. Bianchini and F. Scarselli, “On the complexity of neural network classifiers: A comparison between shallow and deep architectures,” *IEEE Trans Neural Netw Learn Syst*, vol. 25, no. 8, pp. 1553–1565, 2014, doi: 10.1109/TNNLS.2013.2293637.

11

[2] X. Hu, W. Liu, J. Bian, and J. Pei, “Measuring Model Complexity of Neural Networks with Curve Activation Functions,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2020, pp. 1521–1531. doi: 10.1145/3394486.3403203.

[3] S. S. Du and J. D. Lee, “On the Power of Over-parametrization in Neural Networks with Quadratic Activation,” 2018.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Commun ACM*, vol. 60, no. 6, pp. 84–90, Jun. 2017, doi: 10.1145/3065386.

[5] A. R. Barron, “Statistical properties of artificial neural networks,” in *Proceedings of the IEEE Conference on Decision and Control*, 1989, vol. 1, pp. 280–285. doi: 10.1109/cdc.1989.70117.

[6] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, “Activation Functions: Comparison of trends in Practice and Research for Deep Learning,” Nov. 2018, [Online]. Available: <http://arxiv.org/abs/1811.03378>

[7] R. Vidal, J. Bruna, R. Giryes, and S. Soatto, “Mathematics of Deep Learning,” Dec. 2017, [Online]. Available: <http://arxiv.org/abs/1712.04741>

[8] A. R. Barron, M. Li, and L. Valiant, “Approximation and Estimation Bounds for Artificial Neural Networks,” 1994.

[9] G. E. Dahl, T. N. Sainath, and G. E. Hinton, “Improving deep neural networks for LVCSR using rectified linear units and dropout,” in *ICASSP, IEEE International Conference on Acoustics, Speech, and Signal Processing - Proceedings*, Oct. 2013, pp. 8609–8613. doi: 10.1109/ICASSP.2013.6639346.

[10] E. Mjolsness, C. D. Garrett, and W. L. Miranker, “Multiscale Optimization in Neural Nets,” *IEEE Trans Neural Netw*, vol. 2, no. 2, pp. 263–274, 1991, doi: 10.1109/72.80337.

Collecting relevant articles to the topic which explain the research questions you have produced is a particularly important task when selecting articles. Using a database that is from a credible source is one of the most important things when it comes to choosing articles when conducting research. By using credible sources from databases like ACM, IEEE, and google scholar, I was able to choose the article I believe are a good fit for my research. Our university’s library provides a subscription to ACM, IEEE, ScienceDirect, and other credible sources to help ensure people have access to articles in the research, especially for the IT and Data Science fields.

Using Boolean operators like AND as well as OR, help is filtering out the general search of articles to a more precise list of articles. The usage of quotations on specific words like “Activation function,” “neural networks,” and “mathematics” helped in reducing the number of articles from my initial search.

Also, when choosing an article, focusing on the timeframe is important. However, make sure not to miss important articles from the past and avoid recency bias. Yes, it is good to look at recent articles, since the IT industry is changing rapidly, but important articles from the past that helped build the foundation for this topic should not be discarded because they were published in the past.

The validity of an article is also a major component in choosing an article for research. The validity of an article can be internal validity or external validity. In my research, most of the validities are external validity because, we focused on reproducibility, generalizability, and randomization of the calculations when discussing how activation functions are explained from a mathematical point of view.

1. **Using the best practices stated in the article, “Writing a literature review,” write a systematic literature review that provides a cohesive and analytical summary for five of the articles (about 250-300 words per article). Use the “cite while you write” feature of the citation management software to appropriately cite the articles in the ACM or IEEE format. (25 points)**

Understanding the mathematical ways of artificial neural networks when it comes to activation functions is important. The mathematical ways may vary from the calculus aspect to the geometrical and statistical properties of artificial neural networks. The statistical aspect of neural networks explains the network approximation, the convergence of neural networks, and the performance criteria of various activation functions. Focusing on the statistical aspects of neural networks, “the statistical risk of sequences of network estimators converges to zero, as the size of the training sample increases”[1]. Neural networks can be either a nonlinear transformation of linear combinations of multiple variables or a linear combination of nonlinear transformations with fewer variables. Different classes of countless number layers of networks have different implementation functions with different fixed sigmoidal functions and polynomials. Because of the various functions used, the layer of networks will define how fast or slow the convergence rate is for different networks. The performance is dependent on the choices of loss functions and the estimation criteria. There is no best loss function however based on different research the zero-one loss function seems to be the “best” compared to other loss functions, but different loss functions have different uses depending on the requirements. Regulating the complexity of the networks is also something of importance. “Network convergence theorem: assume that several of approximation tends to zero. If the range for every function is in a fixed interval of length b, the definition of complexity regularized estimator, then the statistical risk of the network estimator converges to zero at rate bounded” [1].

With the convergence of neural networks, “finding an optimal neural network architecture is vital for both accuracy and performance” [2]. Based on the article *Activation Functions: Comparison of trends in Practice and Research for Deep Learning*, normal distribution was used in the creation of arrays containing 1000 x-values with a mean of zero and various standard deviations between different arrays. When comes to activation functions, it is calculated based on the sum of the activations that neurons are in linear functions. “Loss function is a function defined on a neural network and a training set. A loss function provides a way of measuring how far off the predictions of the neural network are on a training set” [2]. Comparing the three most common activation functions, ReLU, Sigmoid, and Tanh, studies have shown that ReLU is better. However, based on this article, Sigmoid had a performance rate of 10.1% higher compared to ReLU. “An optimizer is the algorithm used to minimize the loss function. The most common optimizer is Gradient Descent or Stochastic Gradient Descent (SGD). Gradient descent works by first calculating the derivative of the loss function and then changing the weights and biases in the network in the direction of the calculated gradient reducing the loss function in a series of steps” [2]. By pairing a function and an optimizer together, ReLU is the best performing when paired with an SGD optimizer. To say one is better than the other activation functions, understanding the complexity of the functions is an important factor.

To reduce the complexity of the functions, it is claimed that deep learning is a solution. “deep neural networks can implement functions with higher complexity than shallow ones, when using the same number of resources” [3]. As stated previously, having a multilayer network system will produce a more accurate and precise result. “The network implements a function by activating each neuron, starting from the first layer, which actually “passes” the input, to the last layer, which returns the calculated function,” [3]. Other than the multilayer neural networks, the article, *On the complexity of neural network classifiers: A comparison between shallow and deep architectures,* talk about Betti Numbers and Pfaffian functions. “Formally, the kth Betti number, bk(S), is defined as the rank of the kth homology group of the space S” [3]. The bounds of activation functions are the limitation, therefore understanding the upper and lower bounds is essential. “an upper bound on the sum of the Betti numbers of the positive set realized by three-layer networks with arctan(·) activation function”[3]. This article used multiple theorems for the upper bounds of the activation functions. For lower bounds “In the following, we present some results on lower bounds of B(SN), for both deep and shallow networks. In particular, we study the growth of B(SN) when the number of layers is progressively increased” [3]. “In fact, the functions implemented by neural networks belong to a subclass of the Pfaffian functions, so that tighter bounds could be more easily obtainable” [3]. Since deep learning increases the accuracy of activation functions, improving deep neural networks is important.

Until recently, the majority of state-of-the-art speech recognition systems were based on hidden Markov models (HMMs) that employed Gaussian mixtures to simulate the HMM emission distributions. However, demonstrated that hybrid acoustic models that substituted Gaussian mixture models (GMMs) with pre-trained, deep neural networks (DNNs) could significantly enhance performance on a small-scale phone detection challenge, which was then extended to a broad vocabulary voice search job in. Since then, various researchers have made huge advances in long vocabulary continuous speech recognition (LVCSR) tests by utilizing deep neural network acoustic c models. Wide and deep neural networks are subject to overfitting even with unsupervised pre-training and big training sets. Although dropout is simple to implement in minibatch stochastic gradient descent, the optimal technique to incorporate it into 2nd order optimization methods remains an outstanding research subject. Although we employed minibatch stochastic gradient descent for all frame-level training, we elected to use the Hessian-free optimizer for the whole sequence training in this study since it had significant parallelization advantages over the SGD sequence training system, we had available. However, one unfavorable side effect of this decision is the undesired interplay between dropout and HF. To locate each search direction, HF repeatedly optimizes a local quadratic approximation to the objective function using the conjugate gradient (CG) technique. Because it makes such extensive use of curvature and gradient information generated from big batches, CG relies on it. “Extracting bottleneck features using deep nets trained with dropout would obviate most of the need for sequence training of the nets, but would abandon the goal of end-to-end training using a sequence level criterion. Nevertheless, it is a promising alternative we hope to explore” [4].

Continuing on the model complexity of neural networks, “Deep neural networks have gained great popularity in tackling various real-world applications, such as machine translation”[5]. Model structures, such as breadth, layer depth, and layer type, can have a significant impact on model complexity. The power of layer width in shallow neural networks has been studied for decades, with the universal approximation theorem stating that a single-layer feedforward network with a finite number of neurons may approximate any continuous function with some basic assumptions. Some further research strengthens this theorem. The universal approximation theorem, on the other hand, allows the layer width to be exponentially huge. The universal approximation theorem is extended to deep networks with restricted layer breadth. Deep models have recently been empirically proven to be more successful than shallow models. A number of research has been conducted to investigate the benefits of the deep architecture. A deep model's function with piecewise linear activation functions is piecewise linear and has a limited number of linear regions. The number of linear zones in such a model is typically used to estimate the model's nonlinearity or complexity. Motivated by this, we create a piecewise linear approximation of the target model using curve activation functions, then utilize the number of linear regions of the approximation model as a reflection of the target model’s complexity.

1. **Based on your literature review, identify gaps in previous research that your research may aim to address. Identify at least three research questions from your literature review. Write a succinct problem/purpose statement using your research questions (200-300 words). The problem statement should address the following:**
   1. **What would be the impact of your topic and research questions? (Is your topic worth researching?)**
   2. **What gaps will your research address?**
      1. ***points)***
2. What are the ways to optimize neural networks for more precise and accurate results?
3. What are the complexities of neural networks in terms of shallow and deep learning?
4. What are the types of functions used on different activation functions of neural networks?

The purpose of this research is to understand the mathematical aspects of the activation functions of neural networks. Activation functions are used in various places like image recognition. Having an accurate result is an important task as well have optimized the time it takes for the task to be completed. Since we have multiple activation functions, all of them will have individualized mathematical equations with different linear equations and nonlinear transformations. Understanding the use of loss functions and what loss functions are when it comes to the optimization of neural networks. The layering of multiple activation functions (the same activation function layer multiple times) will result in different outcomes because of the accuracy they produce. An understanding of these activation functions will help with knowing when to use them accordingly to obtain the most accurate and precise results. Each function has different upper and lower bounds which are the limitation and will require you to change your function if the task in hand is out of bounds. Therefore, understanding the mathematics behind all or most of the activation functions is important to choose the appropriate function and gaining the best result. That is the purpose of this research proposal on activation functions in neural networks. We present a complexity metric for deep neural networks with curve activation functions in this research. We first propose a piecewise linear framework, the linear approximation neural network (LANN), to both approximate a given DNN model to a necessary approximation degree and reduce the number of generated linear regions. We develop the complexity measure assisted by the upper bound after establishing an upper bound on the number of linear areas created by LANNs. We conducted imperial study to assess the usefulness of the complexity measure, which revealed a positive link between the incidence of overfitting and the rise of model complexity during training.

1. **Identify and write the research objectives (4-5 objectives) using the problem/purpose statement. The research objectives are concise, specific, and measurable. Use Bloom’s Taxonomy to identify appropriate active verbs for your research objectives. *(25 points)***

**The following is an example:**

***My research project will:***

* ***Identify variables that have a statistically significant role in student drop out in online graduate courses.***
* ***Build a predictive model to identify at-risk students in online graduate courses.***
* ***Evaluate machine learning techniques for binary classification based on accuracy.***
* Evaluate the mathematical aspects of activation functions and their use in neural networks.
* Identify the convergence results of the different activation functions
* Identify the complexity of activation functions in terms of the upper and lower bound
* Identify the optimization of neural networks.
* Evaluate the approximation and estimation bounds of activation neural networks

1. **What are the ethical implications of your research? How do you plan to address them? The ethical issues may appear in any phase of your research including data collection, data management, data analysis, conclusions, and dissemination of findings. *(20 points)***

Since this research focuses on the mathematical aspects of activation functions, the ethical implications are minimal. The main ethical implications would be plagiarism and honesty. Plagiarism is the appropriation of another person’s ideas, processes, results, or words without giving appropriate credit. For this research, it is important to give the appropriate credit and have proper citations to give credit where credit is due. Honesty is always an ethical implication that should be focused on in every research. Honesty focuses on reporting correct results, not falsifying or fabricating data, and giving proper credit.

1. **The design science methodology can be used to the design of an IT artifact (such as an interface, a database, or an IT system). How is design science different from other perspectives (worldviews) such as positivist (or post-positivist) and interpretive (such as constructivist). Discuss your answer in the context of ontology, epistemology, methodology, and axiology. Cite appropriate research papers that you used to answer the question. *(15 points)***

“Design science research is used to develop modern technologies for solving problems. Such problems and solutions are often socio-technical in nature, which creates problems for DSR in gaining problem understanding, identifying systemically appropriate solutions, and in effectively evaluating new and innovative solutions” [6]. There are different guidelines when it comes to design science, around seven guidelines. Here are the following guidelines and their descriptions.

**Guideline 1: Design as an Artifact**

Design science has to produce a viable artifact in a form of a construct, model, method, or instantiation.

**Guideline 2: Problem Relevance**

Design science is important in the development of technology-based solutions to important and relevant business problems.

**Guideline 3: Design Evaluation**

The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.

**Guideline 4: Research Contributions**

Design Science must provide a clear and verifiable contribution in the areas of design artifacts, design foundation, and design methodologies.

**Guideline 5: Research Rigor**

Design Science relies on the application of rigorous methods in both the construction and evaluation of the design artifact.

**Guideline 6: Design as a Search Process**

Design Science requires utilizing available means to reach desired ends while satisfying the laws in the problem environment.

**Guideline 7: Communication of Research**

Design Science has to be presented effectively to both technology-oriented and management-oriented audiences.

When it comes to worldviews, there are diverse types of worldviews. The types of worldviews are postpositivist, constructive, pragmatic, and transformative worldviews.

The Postpositivist worldview is usually quantitative and focuses on the cause and effect of a problem. Since it is a cause-and-effect type of worldview, it begins with different theories and the collection of data to either support or refutes the theories. The theories are generated through deduction. Objectivity is essential when it comes to postpositivist views because researchers must examine methods and conclusions for bias and maintain the standards of reliability and validity in quantitative research. It is a type of worldview that is based on the knowledge of data, evidence, and rational considerations. Post-positivism is usually used in fields like computer science, data science, and so on.

The constructive worldview is usually qualitative with open-ended questions. Social constructivists believe individuals seek understanding of the world they live in, meaning individuals look at the world through their lens of experiences resulting in many varied and complex views. In a constructive worldview, theories are generated through induction.

Ontology is a subject area that shows the relationship and properties of distinct categories. By this definition, ontology is more postpositivist since it focuses on the cause-and-effect of distinct categories and how each of them is related to the other.

Epistemology is a theory/study that focuses on people’s beliefs and opinions. Using the definition, epistemology is similar to constructive worldview because it believes in people’s opinions and focuses on how individuals have their own experiences and opinions based on the lives they have lived.

The methodology is used in a particular study. Design Science is focused on particular studies and has diverse ways to learn more about that particular topic.

1. **What is validity in research? How is validity different from reliability and accuracy? Differentiate between external and internal validity using examples. *(15 points)***

“Validity is a concept that is used about measures, which is a generic term for measurement devices (e.g., tests, questionnaires, interview schedules, or personality scales).”[6], [7]. Validity is a significant component to authenticate the quality of research. Validity can be internal and external validity.

Internal validity depends on the procedures of the study and how rigorously it is performed. Internal validity is a study that established a trustworthy cause-and-effect relationship between a treatment and an outcome. Internal validity can be improved using randomization and random selection of participants, blinding, and study protocols. On the other hand, internal validity can be threatened by confounding variables, biased sampling, and instrumentation errors.

External validity refers to how well the outcome of a study can be expected to apply to other settings. External validity focuses on the generalizability of the findings. This includes things like do the findings apply to other people, settings, situations, and periods. External validity can be improved using inclusion and exclusion criteria, field experiments, and replication. At the same time, external validity can be threatened by selection biased, sample features.

Reliability is similar to precision in that it focuses on how the same/comparable results are being hit. These results might not be accurate to the correct/actual results. However, if they are all producing the same results, even if it is incorrect, it is a reliable result. Accuracy, on the other hand, is if the test hits the correct result even if it is just one.

Validity is more fulfilled when both accuracy and reliability are fulfilled together. If the test results a correct result and can be achieved on multiple tests, then it is a valid test.

**Reference**

[1] A. R. Barron, “Statistical properties of artificial neural networks,” in *Proceedings of the IEEE Conference on Decision and Control*, 1989, vol. 1, pp. 280–285. doi: 10.1109/cdc.1989.70117.

[2] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, “Activation Functions: Comparison of trends in Practice and Research for Deep Learning,” Nov. 2018, [Online]. Available: http://arxiv.org/abs/1811.03378

[3] M. Bianchini and F. Scarselli, “On the complexity of neural network classifiers: A comparison between shallow and deep architectures,” *IEEE Trans Neural Netw Learn Syst*, vol. 25, no. 8, pp. 1553–1565, 2014, doi: 10.1109/TNNLS.2013.2293637.

[4] G. E. Dahl, T. N. Sainath, and G. E. Hinton, “Improving deep neural networks for LVCSR using rectified linear units and dropout,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, Oct. 2013, pp. 8609–8613. doi: 10.1109/ICASSP.2013.6639346.

[5] X. Hu, W. Liu, J. Bian, and J. Pei, “Measuring Model Complexity of Neural Networks with Curve Activation Functions,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2020, pp. 1521–1531. doi: 10.1145/3394486.3403203.

[6] Vijay. Vaishanvi and ACM Digital Library., *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology.* ACM, 2009.

[7] “Understanding Research Methods”.

[8] S. S. Du and J. D. Lee, “On the Power of Over-parametrization in Neural Networks with Quadratic Activation,” 2018.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Commun ACM*, vol. 60, no. 6, pp. 84–90, Jun. 2017, doi: 10.1145/3065386.

[10] R. Vidal, J. Bruna, R. Giryes, and S. Soatto, “Mathematics of Deep Learning,” Dec. 2017, [Online]. Available: http://arxiv.org/abs/1712.04741

[11] A. R. Barron, M. Li, and L. Valiant, “Approximation and Estimation Bounds for Artificial Neural Networks,” 1994.

[12] E. Mjolsness, C. D. Garrett, and W. L. Miranker, “Multiscale Optimization in Neural Nets,” *IEEE Trans Neural Netw*, vol. 2, no. 2, pp. 263–274, 1991, doi: 10.1109/72.80337.

Submission: Submit a Word document containing your answers in D2L. Name your document *CIS600\_Individual\_Assignment\_YourLastName.docx*. Submit the mp4 file for question 2. Name your video *CIS600\_Individual\_Assignment\_YourLastName.mp4.*