# Formal Project Proposal template

Due 27 June 2023

|  |  |
| --- | --- |
| **Name:** | Thomas Grubb |
| **Programme:** | Data Science - MSc |
| **Student ID:** | 09029648 |
| **Project Title:** | Machine Learning Identification of the Haast Tokoeka Kiwi Bird from Acoustic Readings" |
| **Problem you are seeking to address:** | The Haast tokoeka kiwi bird (HTK) stands as one of New Zealand's rarest avian species, with an estimated population of merely 400 individuals in 2015. Indigenous to the challenging mountainous terrains of AAST, Fiordland, and Rakiura in the South Island of New Zealand (Aotearoa), the Haast tokoeka kiwi's elusive nature poses a unique challenge for conservationists. In an effort to safeguard the species, preserving genetic diversity within the existing population is paramount.  Given the difficulty of navigating the bird's habitat, acoustic recorders have become invaluable tools in the identification of the size of current populations and the discovery of new breeding pairs. These devices capture audio data, offering a potential means to identify the Haast tokoeka kiwi bird within the vast and challenging landscapes. Accurate identification is pivotal for determining population sizes and confirming the existence of new populations, crucial aspects in the species preservation strategy.  In collaboration with the Department of Conservation, I am eager to contribute to the analysis of this data. The objective is to employ machine learning methods to enhance the accuracy of identifying the Haast tokoeka kiwi bird in these sound recordings. This initiative aims to provide conservationists with a robust and efficient tool for monitoring and safeguarding this critically endangered species. |
| **Aim and scope:** | My major aim is to create an algorithm that can reliably identify the HTK from sound recordings. I hope to investigate and review the effectiveness of both front end and back-end techniques used in the classification of sound data. If successful I may extend the project into identifying numbers of birds on the recordings and other animals/birds in the local area. |
| **Project objectives:** | 1. *Create/obtain a data set of recordings that contains the* HTK and other birds and animals in the local area from a set of recordings provided by the department of conservation and using the Xeno-Canto database to help supplement the data set. 2. *Use appropriate methods to help clean the data of noise by use of noise removal algorithms.* 3. *Investigate the appropriateness of various methods of feature extraction for sound data :*  * *Mel Frequency Cepstral Coefficients* * *Non-negative matrix factorisation*  1. *Use and analyse the accuracy of the results of using various machine learning methods to such as:*  * *K Nearest neighbour* * *Neural networks* * *Random forest* * *Regression (for classification)* * *Gradient boosting* * *Support vector machines* |
| **Expected project outcomes:** | *I will produce several machine learning models that will take a sound recording and identify whether a HTK is present, documenting the accuracy of each model. I hope to achieve a level of accuracy that is sufficient to be used by conservationists in their work.*  *I will conduct research into feature extraction of sound data in order to apply this to my model. I will also research models that have already been used to identify bird calls and analyse their effectiveness.* |
| **Brief review of relevant literature:** | *Cakmamk et al. used spectrographic analysis to classify Rose-crested pipits in their paper “Interactive Classification Using Spectrograms and Audio Glyphs”, this included extracting the spectral centroid, spectral bandwidth and the spectral roll-off. They displayed their results in audio glyphs which they used to classify the birds.*  *Ludena-Choez et al. proposes the use of non-negative spectrogram decomposition as a more effective way of extracting features from acoustic data. In their study they found that this outperformed more traditional approaches such as Mel frequency cepstral coefficients.*  *Nani et al. achieved 97% accuracy in their deep learning classifier on a bird dataset using a convolutional neural network (CNN) . They concluded that the best standalone CNNs are VGG16 and VGG19 although they stated that a big draw back of CNNs is that they require a very large data set which will not be provided for my project.*  *Mortimer and Greene discuss in their paper the uncertainty of classification of bird song by professors with audio recordings of the Tui bird only being agreed upon 50% of the time. This highlights the need for robust scrutiny of the data I am provided.*  *Ghoraani and Krishnan used a three stage process to extract features from an audio signal which comprises of constructing a time-frequency matrix they found that this method gave more accurate results compared with the results gained from features extracted by spectrogram analysis.* |
| **[optional] References:** | Anon. (no date) *Save the Kiwi* [online]. Available from: <https://savethekiwi.nz/about-kiwi/kiwi-species/tokoeka/haast-tokoeka/> [Accessed 29 February 2024a].  Anon. (no date) [online]. Available from: <https://nzbirdsonline.org.nz/location-search?field_location_term_id=149&field_location_term_value=Westland> [Accessed 1 March 2024b].  Anon. (no date) [online]. Available from: <https://www.doc.govt.nz/nature/native-animals/birds/birds-a-z/kiwi/tokoeka/> [Accessed 29 February 2024c].  Anon. (no date) [online]. Available from: <https://xeno-canto.org/> [Accessed 29 February 2024d].  Cakmak, E., Schlegel, U., Miller, M., Buchmüller, J., Jentner, W. and Keim, D.A. (2018) *Interactive Classification Using Spectrograms and Audio Glyphs*. In: *2018 IEEE Conference on Visual Analytics Science and Technology (VAST)* [online]2018 IEEE Conference on Visual Analytics Science and Technology (VAST). pp. 110–111. Available from: <https://ieeexplore.ieee.org/document/8802500> [Accessed 29 February 2024].  Ghoraani, B. and Krishnan, S. (2011) Time–Frequency Matrix Feature Extraction and Classification of Environmental Audio Signals. *IEEE Transactions on Audio, Speech, and Language Processing* [online]IEEE Transactions on Audio, Speech, and Language Processing. 19 (7), pp. 2197–2209.  Ludeña-Choez, J., Quispe-Soncco, R. and Gallardo-Antolín, A. (2017) Bird sound spectrogram decomposition through Non-Negative Matrix Factorization for the acoustic classification of bird species. [online]. p. e0179403.  Marsland, S., Priyadarshani, N., Juodakis, J. and Castro, I. (2019) AviaNZ: A future-proofed program for annotation and recognition of animal sounds in long-time field recordings. *Methods in Ecology and Evolution* [online]. 10 (8), pp. 1189–1195.  McLoughlin, I., Xie, Z., Song, Y., Phan, H. and Palaniappan, R. (2020) Time–Frequency Feature Fusion for Noise Robust Audio Event Classification. *Circuits, Systems, and Signal Processing* [online]. 39 (3), pp. 1672–1687.  Nanni, L., Maguolo, G., Brahnam, S. and Paci, M. (2021) An Ensemble of Convolutional Neural Networks for Audio Classification. *Applied Sciences* [online]. 11 (13), pp. 5796-.  Robertson, H.A. and De Monchy, P.J.M. (2012) Varied success from the landscape-scale management of kiwi *Apteryx* spp. in five sanctuaries in New Zealand. *Bird Conservation International* [online]. 22 (4), pp. 429–444.  Tsalera, E. 1, Papadakis, A. 2, Samarakou, M. 1 1 D. of I., Computer Engineering, S. of E., etsalera@uniwa.gr (E.T.), Electrical, marsam@uniwa gr (M S.) 2 D. of, Electronic Engineering Educators, S. of P. and Technological Education (ASPETE), 15122 Athens (2021) Comparison of Pre-Trained CNNs for Audio Classification Using Transfer Learning. [online]. p. 72.  Zhang, X.-Y. and He, Q.-H. (2015) Time–frequency audio feature extraction based on tensor representation of sparse coding. *Electronics Letters* [online]. 51 (2), pp. 131–132. |
| **Testing and Evaluation:** | *I will be using traditional measures for identifying the effectiveness of my classification model. These will include:*   * *Confusion matrix* * *ROC-AUC* * *F1 Score* * *The ratio of correctly predicted instances against the total instances (accuracy)* * *The ratio of correctly predicted positive instances against the total instances of a positive result (precision)* * *The ratio of correctly predicted positive instances against the total instances recall)* |
| **Ethical considerations:** | *The data I will be using will be sourced via a third party. This will require adherence to any confidentiality agreements they may wish to issue me, I will require the ability to publish my work and to be able to change any data that I am gifted. All data processing will be documented, and code will be replicable. Similarly, the online database xeno-canto that I will be using to supplement my data set requires that I do not use any data stored for financial gain.* |
| **Project plan:** |  |
| **Signature:** |  |