**Task 1**

Introduction

* The objective of the project is to determine which uncultivated lands in a specified area in Malta have the potential to be cultivated,
* By making use of sentinel imagery to obtain visual changes in agricultural land in Malta across different periods.
* Actual data gathered from the Agriculture and Rural Payments Agency (ARPA) about which lands are cultivated and uncultivated.
* Making use of different algorithms, both supervised and unsupervised classifiers, to compare which would obtain better results.

Completed work

* During Summer, I carried out research on relative papers which focus on remote sensing and agricultural studies. I gathered these papers to be used for the Literature Review later.
* Weekly mentor meetings are being held to discuss how to move forward in building up my project.
* ARPA were contacted to issue me with ground truth data on which land was found to be cultivated and which was uncultivated from their on-the-spot inspections. A meeting was set with the Senior Manager where I went on their premises to discuss my project and what data they can provide me with.
* In the meantime, I set two different polygons for cultivated and uncultivated lands classes and a study area class on QGIS to train on and I followed a tutorial which carried out a Random Forest (RF) Classification based in Seville on both a single satellite image and on multiple images. Hence, I also downloaded 3 different satellite images to carry out the RF.
* After some time, ARPA forwarded me a complete dataset as parcels from their inspections, including agricultural land in all around Malta which is cultivated and that which is not.
* This data was transferred to QGIS, a geographic information system application that supports viewing, editing, and analysis of geospatial data. A google earth map layer was placed underneath, and ARPA’s dataset was placed over, dictating which lands go under the specified class. The classes were then named as *Arable*, *Non-Cultivated*, and *Other*. A new study area covering a larger area was formed by drawing a polygon around the area I want to focus for my study, that being the outskirts of Siggiewi, Qrendi and Haz-Zebbug. The remaining dataset from the parcels which fell outside of the study area were discarded.
* A time phase was decided for the study to be focused on to determine the changes throughout. Hence, cloud-free satellite imagery was downloaded from the Copernicus Hub platform: 12 satellite images from 2015-2017 and another set of 12 from 2019-2021, by alternating months.
* Random Forest, a Supervised classification, was again carried out on SNAP (Sentinel Application Platform), an ideal architecture for Earth Observation processing and analysis, by using both sets of imageries. A final, masked image for both sets was generated as a GeoTiff displaying the predicted cultivated and uncultivated lands along with the other features such as buildings, streets, and quarries.
* The GeoTiff images were transferred to QGIS, and a Postprocessing for the Landcover Change was done from the Semi-Automatic Classification Plugin (SCP). However, the images being masked on SNAP didn’t return the same Extent values for both images, resulting in SPC postprocessing to return an empty confusion matrix, with area in degree2 as opposed to m2. Hence, the GeoTiffs were once again masked on QGIS, resulting in both having the same Extent value and proper results after running the SPC postprocessing for land cover change, by inputting the 2015-2017 mask as the reference classification, and the 2019-2021 mask as the new classification.
* In the meantime, during the prototype process, I focused on writing up my Literature Review, making use of the papers I found during the Summer and afterwards. A comparison was done to see which algorithms perform best to give better results. The literature review was split into different sections discussing the different types of algorithms possible to be carried out for the purpose of my study and how they work. The first part of the literature review focuses on the changes in Malta’s agriculture throughout the years. First study was published in 2006 by Markou et al., followed by another study by NSO in 2016 which covers the period between 2005 and 2013, a Farm Structure Survey carried out in 2016 by the European Commission in 2021, and the latest Census of Agriculture 2020 by NSO in 2022.

Work to be done

* Using PyCharm, an integrated development environment used in computer programming, specifically for the Python programming language (Wikipedia, 2022), a K-means classification will be applied to two of the downloaded satellite images, one from each set of periods. To do so, each band must be extracted from the resampled images in SNAP by opening each band’s image window and choosing only the band related to it from the ‘Spatial subset from view’ tool.
* After each band has been extracted (12 in total), these will then be opened in PyCharm to read the data. Firstly, a library called EarthPy must be installed, which is a python package that makes it easier to plot and work with spatial raster and vector data using open-source tools (Earth Lab, 2021). From then, I would need to visualize the data form the inputted bands and apply an RGB composite Image. A data distribution of Bands will be outputted, displaying a histogram for each band and its colour.
* I would need to write up the methodology section, discussing what system of methods were used for my study and how I went about obtaining the results. I would need to look at other relatable studies and compare their methodologies to see which would comply best with my purpose.
* If need be, I would make use of further algorithms for the evaluation phase, to be able to analyze results in different approaches and compare which would be most suited for my study.
* Finally, I would focus on writing up the introduction, conclusion, and abstract sections of the paper.

Conclusion

**Task 2**

Being a relatively new area of study, findings and research are still scarce, and the number of sources focusing on remote sensing in agriculture was limited. Particularly taking into consideration that my study focuses on Malta, where no other forms of related studies were found. The majority of the studies found are based on large map areas as compared to my study area which is quite small, using specific methods and approaches which might not be ideal for my purpose, and this might lead to inconsistent results than that of the found studies. Even though the concept map is in a spider mapping format, which is typically the easiest type of concept map to make and interpret, it might be visually intimidating or chaotic when used to investigate a particular subject. It might also seem like a vague concept map due to being limited with keywords.

References:

Abdi, A, M, 2020, Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data, GIScience & Remote Sensing, 57:1, 1-20, DOI: 10.1080/15481603.2019.1650447

Adugna, T, Xu, W, Fan, J 2022, Comparison of Random Forest and Support Vector Machine Classifiers for Regional Land Cover Mapping Using Coarse Resolution FY-3C Images, Remote Sens. 2022, 14(3), 574; <https://doi.org/10.3390/rs14030574>

Atriga Consult, 2018, National Agricultural Policy for the Maltese Islands 2018 – 2028, Available at: <https://meae.gov.mt/en/Public_Consultations/MSDEC/Documents/National%20Agricultural%20Policy%20for%20the%20Maltese%20Islands%202018%20-%202028.pdf>, [Accessed 8 February 2022].

Bai, S, Kolter, J, Z, Koltun, V, 2018, An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling, arXiv:1803.01271v2

Bauckhage, C, 2013, Lecture Notes on k-Means Clustering (I), B-IT, University of Bonn, DOI: 10.13140/2.1.3738.4480

Belgiu, M & Csillik, O, 2018, Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis, Remote Sensing of Environment: 204, pp509-523

Bhaduri, A., 2021. Geographic Information Systems Courses: After 12th, Online, Eligibility, Syllabus, Top Colleges, Syllabus, Scope, Jobs. [online] Collegedunia.com. Available at: <https://collegedunia.com/courses/geographic-information-systems> [Accessed 2 February 2022].

Burrough, P.A., 2001. GIS and geostatistics: essential partners for spatial analysis. In: Environmental and Ecological Statistics. vol. 8. Kluwer Academic Publishers, pp. 361–377

CUOSGwiki, 2022. Random Forest Supervised Classification Using Sentinel-2 Data - CUOSGwiki. [online] Dges.carleton.ca. Available at: <https://dges.carleton.ca/CUOSGwiki/index.php/Random_Forest_Supervised_Classification_Using_Sentinel-2_Data#Sentinel_Application_Platform_.28SNAP.29> [Accessed 2 February 2022].

Cutler, A, Cutler, D, R, Stevens, J, R 2011, Random Forests, Machine Learning 45(1):157-176, Ensemble Machine Learning: Methods and Applications, DOI: 10.1007/978-1-4419-9326-7\_5

ESA, 2019. *Sentinel Application Platform - eo science for society*. [online] eo science for society. Available at: <https://eo4society.esa.int/resources/snap/> [Accessed 2 February 2022].

European Commission, 2021, Statistical Factsheet Malta, Available at: <https://ec.europa.eu/info/sites/default/files/food-farming-fisheries/farming/documents/agri-statistical-factsheet-mt_en.pdf>, [Accessed 2 February 2022].

Ghioca-Robrecht, D. M., C. A. Johnston, and M. G. Tulbure.2008. “Assessing the Use of Multi-season Quick Bird Imagery for Mapping Invasive Species in a Lake Erie Coastal Marsh.” Wetlands 28 (4): 1028–1039. doi:10.1672/08-34.1.

Ghosh, S, Saraf, S, Behera, M, Biradar, C, 2017, Estimating Agricultural Crop Types and Fallow Lands Using Multi-Temporal Sentinel-2A Imageries. Proceedings of the National Academy of Sciences, India Section A: Physical Sciences. 87. 10.1007/s40010-017-0447-5.

Z. Jabs-Soboci ́nska, A. N. Affek, I. Ewiak, and M. D. Nita, 2021, “Mapping mature post-agricultural forests in the polish eastern Carpathians with archival remote sensing data,” Remote Sensing, vol. 13, no. 10, 2021. [Online]. Available: <https://www.mdpi.com/2072-4292/13/10/2018>

Li, X, Lv, X, He, Y, Zhou, B, Deng, J, Qin, A, 2021, Application of Random Forest in Identifying Winter Wheat Using Landsat8 Imagery, Engenharia Agrícola, Jaboticabal, v.41, n.6, p.619-633 Doi: <http://dx.doi.org/10.1590/1809-4430-Eng.Agric.v41n6p619-633/2021>

Markou, M, Kavazis, Staavris, 2006, Agricultural Situation Report – Malta, National Agricultural Policy, DOI: 10.13140/RG.2.1.3062.6965, Available at: <https://www.researchgate.net/publication/299392964_Agricultural_Situation_Report_-_Malta>, [Accessed 8 February 2022].

Morissette, L, Chartier, S, 2013, The k-means clustering technique: General considerations and implementation in Mathematica, Tutorials in Quantitative Methods for Psychology 9(1):15-24, DOI: 10.20982/tqmp.09.1.p015

NSO, 2016, Agriculture and Fisheries 2014, Available at: <https://nso.gov.mt/en/publicatons/Publications_by_Unit/Documents/B3_Environment_Energy_Transport_Agriculture_Statistics/Agriculture_and_Fisheries_2014.pdf>, [Accessed 8 February 2022].

NSO, 2022, Census of Agriculture 2020, Available at: <https://nso.gov.mt/en/News_Releases/Documents/2022/02/News2022_015.pdf>, [Accessed 8 February 2022]

Paramasivam, C, R 2019, Merits and Demerits of GIS and Geostatistical Techniques, GIS and Geostatistical Techniques for Groundwater Science, DOI: 10.1016/B978-0-12-815413-7.00002-X

Pelletier, C, Webb, G, L, Petitjean, F, 2019, Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series, Remote Sens. 2019, 11(5), 523; <https://doi.org/10.3390/rs11050523>

Prishchepov, A. V., V. C. Radeloff, M. Dubinin, and C. Alcantara, 2012, The Effect of Landsat ETM/ETM+ Image Acquisition Dates on the Detection of Agricultural Land Abandonment in Eastern Europe, Remote Sensing of Environment 126:195–209, DOI:10.1016/j.rse.2012.08.017

P. Rao, W. Zhou, N. Bhattarai, A. K. Srivastava, B. Singh,S. Poonia, D. B. Lobell, and M. Jain, “Using sentinel-1, sentinel-2, and planet imagery to map crop type of smallholder farms,”Remote Sensing, vol. 13, no. 10, 2021. [Online]. Available: <https://www.mdpi.com/2072-4292/13/10/1870>

Roy, P, S, Roy, A, 2010, Land Use and Land Cover Change: A Remote Sensing & GIS Perspective, Journal of the Indian Institute of Science, 90(4):489-502

Srivastava, D, Bhambhu, L, 2010, Data classification using support vector machine, Journal of Theoretical and Applied Information Technology 12(1):1-7

Tong, X, Brandt, M, Hiernaux, P, Herrmann, S, Rasmussen, L, V, Rasmussen, K, Tien, F, Tagesson, T, Zhang, W, Fensholt, R, 2020, The forgotten land use class: Mapping of fallow fields across the Sahel using Sentinel-2, Remote Sensing of Environment 239:111598, DOI: 10.1016/j.rse.2019.111598

Wan, R, Mei, S, Wang, J, Liu, M, Yang, F, 2019, Multivariate Temporal Convolutional Network: A Deep Neural Networks Approach for Multivariate Time Series Forecasting, Deep Learning Applications with Practical Measured Results in Electronics Industries, Electronics 2019, 8(8), 876; <https://doi.org/10.3390/electronics8080876>

**Task 3**

Quantitative studies:

Tong et al. (2020) undertook the study of mapping fallow fields across the Sahel using Sentinel-2. The purpose of this study is expected to distinguish between actively cropped fields and fallowed fields within agricultural lands whilst improving our understanding of crop-fallow rotation dynamics by mapping cropped and fallow land patches across the entire Sahel at a 10m spatial resolution. Sentinel-2 satellite imagery was used to disentangle the named agricultural land, facilitated with Google Earth Engine for big data handling. The first 10 m resolution Sahelian fallow field map was created for the baseline year 2017, by developing a remote sensing-driven approach for producing reference data for mapping over broad regions.

Abdi (2019) conducted a study on the efficiency of machine learning algorithms in classifying land cover and land use in a boreal region using Sentinel-2 data. The intent of this study was to compare and examine in terms of classification the performance of four non-parametric algorithms: Support vector machines, random forests, extreme gradient boosting, and deep learning. A complex mixed-use landscape in south-central Sweden with eight land-cover and land-use classifications was chosen as the research region. Sentinel-2 multi-temporal sceneries including spring, summer, fall, and winter conditions were employed to classify the satellite pictures. Each LCLU class received 1477 samples, which were separated into training (70%) and assessment (30%) subsets using stratified random sampling. The accuracy of the algorithm was measured using metrics generated from an error matrix, but the total accuracy was utilized to determine the algorithm hierarchy.

He and Zhao (2019) attempted a study of using Temporal Convolution Networks (TCN) to detect anomalies in time series. The purpose of this study was to utilize TCN to forecast patterns in several time steps after training it on normal sequences. A multivariate Gaussian distribution is used to model prediction errors and produce point anomaly scores. In addition, the performance of a multi-scale feature mixing approach is improved. On three real-world datasets, the method's validity is validated.

**Task 4**

Research Design:

For my study, I collected a specific dataset of ground truth data from the Agriculture and Rural Payments Agency (ARPA), which get inspectors to go on survey inspections to gather which agricultural land is being used and cultivated and which is not. They provided me with this data in agreement with the purpose of my study to be able to conduct my research on whether using Sentinel imagery and machine learning could provide the means of identifying the potential of uncultivated land to be cultivated. This data consists of three parcels declared as No-Minimum (NMM) agricultural activity which consists of land where no ploughing takes place; the declaration is confirmed, and another observed as bare soil; the declaration is not correct. The other cases represent the Arable land class, which also includes no ploughing and bare soil class.

References

EPSG, 2019, *EPSG:32633*, Available at: <https://epsg.io/32633> (Accessed on 2 March 2022).

Earth Lab, 2021, *EarthPy: A Python Package for Earth Data*, Available at: <https://earthpy.readthedocs.io/en/latest/> (Accessed: 3 March 2022).

Wikipedia, 2022, *PyCharm*, Available at: <https://en.wikipedia.org/wiki/PyCharm> (Accessed: 3 March 2022).