Onion irrigation treatment inference using a low-cost hyperspectral scanner

Conference Paper · October 2018

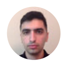
DOI: 10.1117/12.2325500

5 authors, including:


Tiebiao Zhao
University of California, Merced


Haoyu Niu
University of California, Merced


YangQuan Chen
University of California, Merced


Alexander Koumis
University of Southern California

PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Onion irrigation treatment inference using a low-cost hyperspectral scanner

Tiebiao Zhao, Alexander Koumis, Haoyu Niu, Dong Wang, YangQuan Chen

Tiebiao Zhao, Alexander Koumis, Haoyu Niu, Dong Wang, YangQuan Chen, "Onion irrigation treatment inference using a low-cost hyperspectral scanner," Proc. SPIE 10780, Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques and Applications VII, 107800D (23 October 2018); doi: 10.1117/12.2325500

SPIE.

Event: SPIE Asia-Pacific Remote Sensing, 2018, Honolulu, Hawaii, United States

Onion Irrigation Treatment Inference Using A Low-cost Hyperspectral Scanner

Tiebiao Zhao[a], Alexander Koumis[b], Haoyu Niu[a], Dong Wang[c], and YangQuan Chen[a]

[a]University of California, Merced, 5200 Lake Rd, Merced, California, USA
[b]University of Southern California, Los Angeles, California, USA
[c]USDA-ARS Water Management Research Unit, San Joaquin Valley Agricultural Sciences Center, Parlier, California, USA

ABSTRACT

Many studies have shown that hyperspectral measurements can help monitor crop health status, such as water stress, nutrition stress, pest stress, etc. However, applications of hyperspectral cameras or scanners are still very limited in precision agriculture. The resolution of satellite hyperspectral images is too low to provide the information in the desired scale. The resolution of either field spectrometer or aerial hyperspectral cameras is fairly high, but their cost is too high to be afforded by growers. In this study, we are interested in if the low-cost hyperspectral scanner SCIO can serve as a crop monitoring tool to provide crop health information for decision support. In an onion test site, there were three irrigation levels and four types of soil amendment, randomly assigned to 36 plots with three replicates for each treatment combination. Each month, three onion plant samples were collected from the test site and fresh weight, dry weight, root length, shoot length etc. were measured for each plant. Meanwhile, three spectral measurements were made for each leaf of the sample plant using both a field spectrometer and a hyperspectral scanner. We applied dimension reduction methods to extract low-dimension features. Based on the data set of these features and their labels, several classifiers were built to infer the field treatment of onions. Tests on validation dataset (25 percent of the total measurements) showed that this low-cost hyperspectral scanner is a promising tool for crop water stress monitoring, though its performance is worse than the field spectrometer Apogee. The traditional field spectrometer yields the best accuracy as high as above 80%, whereas the best accuracy of SCIO is around 50%.

Keywords: A low-cost hyperspectral scanner, onions, irrigation treatment inference, SCIO

1. INTRODUCTION

Onions are produced and consumed throughout the world. It is worldwide used in different nationalities and cultures during all seasons in a year.[1] California produces the most onion in the US. In 2015, it produced around one third of the total onion crop in the US. It is the only state that can produce spring and summer-harvested onions.[2]

Onions are shallow-rooted crop, and most of the roots were found in the top 0.18 m of soil.[3] This makes it hard for onions to obtain enough soil water. Therefore, lighter and more frequent irrigation are recommended in onion cultural practices.[4] On the other hand, experiments[5] showed that water stress in any growing stages causes reduction in the yield.

To optimize irrigation schedule, it is necessary to have accurate and reliable water stress monitoring methods. Many studies have been published on water stress detection using remote sensing, a real-time and nondestructive method.[6] Near-infrared cameras were used to detected water stress of almond trees,[7–12] where new types of

Further author information: (Send correspondence to YangQuan Chen)
Tiebiao Zhao: E-mail: tzhao3@ucmerced.edu
Alexander Koumis: E-mail: koumis@usc.edu
Haoyu Niu: E-mail: hniu2@ucmerced.edu
Dong Wang: E-mail: dong.wang@ars.usda.gov
YangQuan Chen: E-mail: ychen53@ucmerced.edu

Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques and Applications VII, edited by Allen M. Larar, Makoto Suzuki, Jianyu Wang, Proc. of SPIE Vol. 10780, 107800D
© 2018 SPIE · CCC code: 0277-786X/18/$18 · doi: 10.1117/12.2325500

spectral features were proposed to predict stem water potential. Hyperspectral sensors were also applied for water stress monitoring in apple trees,[13] vineyard,[14,15] etc. In onions, studies of remote sensing are conducted on yield and biomass prediction using the spectroradiometer,[16] biomass monitoring using unmanned aerial vehicles and RGB cameras,[17] detecting diseased onion tissues,[18] quality inspections.[19,20] However, to our best knowledge, nobody has studied irrigation treatment inference using the hyperspectral sensor. Furthermore, SCIO, a low-cost portable, light hyperspectral scanner is evaluated to infer irrigation treatments in onions for the first time.

2. MATERIAL AND METHODS

2.1 Onion Test Site

The test field is in the USDA-ARS, San Joaquin Valley Agricultural Sciences Center(36.594N, 119.512W), Parlier, California. Since 2016, an onion test filed has been set up for research of biomass soil amendments and deficit irrigation. There are three irrigation treatment levels, High, Medium and Low, and four soil amendments, Biochar, Check, Biochar+Compost, and Biochar+Compost+Sulfur. There are three replicate plots for each treatment combination.

2.2 Hyperspectral Scanner

Most recently, a light and small hyperspectral scanner called SCIO (Consumer Physics, Israel) was released in the market. As a complete system, it includes a spectrometer, a light source and optimized algorithms in the cloud. The SCIO spectrometer works in NIR at wavelengths of 700-1100 nm. It is so small that it could even integrated in the smart phone.[21] This system is a low-cost module, with the price less than $300. As a reference, we also used a traditional handhold hyperspectral scanner Apogee PS100 (Apogee, USA). Its wavelength sensitivity is from 350 nm to 1150 nm, with the spectral resolution of 1 nm and digital resolution of 16 bit.

2.3 Field Measurement Collection

During the growing season, onions under different treatment were sampled once a month. The field hyperspectral measurements were coordinated with these physiological measurements including shoot length, root length, number of leaves, fresh weight, dry weight and bulb diameter. There are three onion samples collected for each plot. For each onion sample, three hyperspectral measurements were made using both Apogee and SCIO at the same time in the field. In sum, we have 81 hyperspectral measurements for 27 onions.

SCIO is an active remote sensing platform and it provides calibration case. SCIO was first calibrated using the white panel in the case. The measurement using SCIO requires the distance between leaves and the scanner as small as possible to minimize the disturbance of sun light source. On the other hand, Apogee is a passive remote sensing platform. We first took the measurements of white panel and dark panel for calibration before measured the leaves.

2.4 Principle Component Analysis

Both Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are methods used for reducing dimentionality of a dataset to increase accuracy, speed up processing time, and aide visualization.

PCA is a linear transformation that rotates the axes of the data along the direction that maximizes its variance, allowing data to be projected onto a lower dimentional subspace.[22] These new axes, or "loadings", are found by calculating the eigenvectors W of the data's covariance matrix, where X is an $M * N$ matrix representing M samples of size N:

$$X^T X = W\lambda \tag{1}$$

The eigenvalues λ represent how "important" each loading is in transforming the data, or how much variance the projection onto this axis contributes. As the loadings (and λ) are sorted in descending order, W can be truncated to r columns, which can then be used to project data along r dimensions, preserving the dimensions that contribute most to the variance of distribution. W is often obtained with Singular Value Decomposition (SVD) instead of performing the Eigen decomposition of $X^T X$, as it is more computationally efficient.

2.5 Linear Discriminant Analysis

LDA reduces dimensionality of data by finding new axes to project it onto that maximizes the seperability between classes.[23] It does this by maximizing the distance between means of classes relative to some center point for all classes, while minimizing the variance, or scatter, within each category (Equation. 2). In the following equation, C is the number of classes, N_i is the size of class i, μ is the mean all datapoints, μ_i is the mean of class i, and x_j is the jth datapoint in class i:

$$\frac{\sum_{i=1}^{C} N_i(\mu_i - \mu)(\mu_i - \mu)^T}{\sum_{i=1}^{C}\sum_{j=1}^{N_i}(x_j - \mu_i)(x_j - \mu_i)^T} \tag{2}$$

The optimized solution contains eigenvectors is descending order of their eigenvalues, which can be used to reduce the dataset similar to PCA. Optimizing for both within and between-class scatter is important because only maximizing distance between means could lead to scenarios where the variance is high along the axis with large mean distances, increasing the chance that there are points from different classes overlapping. Minimizing the variance ensures data from each class is grouped tightly along the new axis, increasing seperability.

2.6 Multi-layer Perceptron Classifier

Single perceptrons, or artificial neurons, are nodes with a number of weighted data inputs, a bias input, and an output.[24] The weighted inputs are summed up and fed through an activation function, used as a threshold for when the node should fire, and by how much. They can be used for 2-class classification or regression problems, using either a step function activation function, or nonlinear activation function (e.g. tanh, sigmoid), respectively. The function for a single perceptron is as follows, where s is the activation function, W is the input weights, and b is the bias:

$$f(x) = s(b + Wx) \tag{3}$$

Single perceptrons cannot be used for many complex prediction tasks because they can only predict nonlinear patterns.[25] Multilayer Perceptrons (MLP) overcome this by constructing networks out of multiple perceptrons. An MLP is a supervised learning system consisting of an input layer, N number of hidden layers, and an output layer. Nonlinear activation functions used in MLP's introduce nonlinearities into the model, allowing it to make predictions on complex, nonlinear datasets, such as hyperspectral readings. MLPs are trained using a process called backpropogation, which updates the network's weights with respect to the error between it's current output and the expected result.

3. RESULTS AND DISCUSSION

To prepare the Apogee dataset, each 1675 -dimension reading was loaded into a vector, with readings of the same plant being treated as different data points. SCIO measurements were obtained with the help of smart phone application, and each reading is of 1060 dimensions. For each test iteration, data was split into a 75%/25% train/test size. All data was normalized with Sklearn's normalize() function, while the MLP classifier required normalization with Sklearn's StandardScalar module.[26]

The evaluation stage was broken up into three steps: data preparation, classifier evaluation, and parameter grid search on the best performing classifiers. Classifier performance was ranked by the percentage of correctly predicted labels in the test dataset, averaged over 10 iterations with 75%/25% train/test splits as described above. In the classifier evaluation stage, a number of classifiers from the Sklearn Python package, as well as XGBoost, were ran (with default settings), with and without PCA and LDA dimensionality reduction.

3.1 Results Using PCA Based Classifiers

Figure. 1 shows the performance of several Sklearn classifiers (and XGBoost) ranked against each other in terms of label prediction accuracy for the SCIO dataset (reduced with PCA from 3-99 components). Most classifiers performed very poorly, with only the MLP classifier nearly breaking 50% accuracy, hitting a high score of 48.1% at 90 components, with the rest barely beating random guessing (33% for 3 labels). After MLP was determined

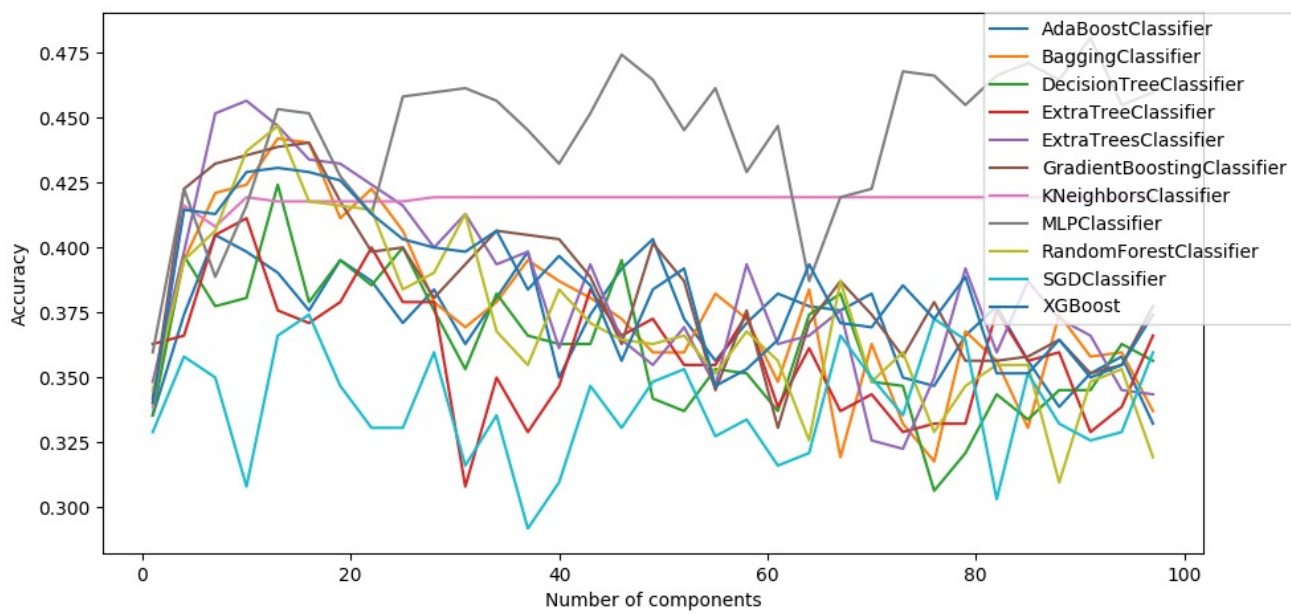


Figure 1: Classifier accuracy comparison on PCA reduced SCIO dataset

to perform the best with PCA reduced data, we performed parameter grid search to optimize it's results, iterating through the parameters listed in 3.1, acheiving a top accuracy of 53.1% (top 5 results and parameters are shown in Table. 1).

- PCA components: 10, 20, 40, 80, 160
- Hidden layer sizes: 25, 50, 100, 200, 400
- Activation: relu, logistic, tanh
- Solver: lbfgs, adam
- Alpha: 0.00001, 0.0001,0.001,0.01,0.1
- Batch size: 200,100,50,25
- Max iterations: 200, 500, 1000

3.2 Results Using LDA Based Classifiers

Each classifier was tested against reduced data with sizes ranging from 3-99 components, as with the PCA reduction in the previous section. This hurt performance relative to PCA reduction, with all classifiers again scoring only slightly above random guessing, with the exception of MLP which scored an average of 38.8% accuracy (see Figure. 2), with a best score of 39.5% at 6 components. As these results were substantially less than the default results of MLP with PCA reduction, parameter grid search was not conducted.

Accuracy	PCA Components	Hidden Layer Sizes	Activation	Solver	Alpha	Batch Size	Max Iterations
0.53064	20	100	tanh	adam	0.001	25	500
0.52903	20	50	relu	lbfgs	0.1	50	500
0.52741	20	200	tanh	adam	1e-05	25	500
0.52741	20	400	relu	adam	0.01	100	500
0.52580	20	50	logistic	adam	1e-05	25	1000

Table 1: Top 5 performing classifiers using PCA and MLP and their grid search parameters

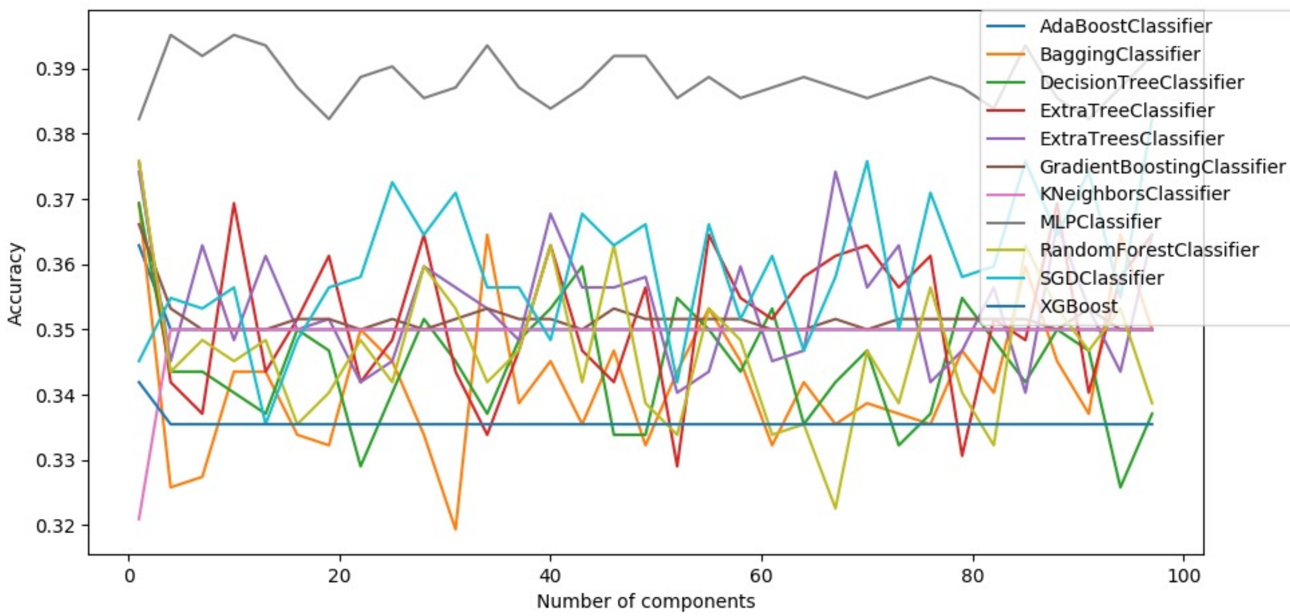


Figure 2: Classifier accuracy comparison on LDA reduced SCIO dataset

3.3 Results Using MLP

Sklearn's MLP implementation was also tested with the raw, unreduced dataset, but performed substantially worse than when it was tested against PCA and even LDA-reduced data. MLP predicted the correct label 36.6% of the time with the unreduced data, essentially randomly picking the result. This demonstrated that dimensionality reduction is necessary for MLP to provide any benefit for this dataset. On the unreduced data, Sklearn's Extra Trees implementation performed the best, predicting the correct label 43.1% of the time.

4. CONCLUSION

Data from the SCIO sensor produces worse classification results as opposed to the Apogee sensor. Using the Apogee data instead of SCIO increased MLP's correct label prediction rate on unreduced data from 36.6% to 71.8%. Similar improvements are seen with dimensionally reduced data, with the MLP classifier configured with default parameters improving its correct prediction rate from 48.1% (with 90 components) to 77.4% (with 30 components) for PCA, and 39.5% (with 6 components) to 58.5% (with 42 components) for LDA. The max results yielded from MLP parameter grid search with PCA-reduced data also increased from 53.1% (with parameters listed in section 3.1) to 83.1%.

In the future, we will test Apogee with more measurements in different growing stages of onions to make sure it is robust to onions' growths. As for SCIO, we will explore new machine learning algorithms to see if better algorithms can help improve prediction accuracy.

ACKNOWLEDGMENTS

Thanks go to Allan Murillo, Zambrzuski, Stella, Christopher Currier, Dong Sen Yan for helping collecting field measurements. Thanks for the support of NVIDIA Corporation with the donation of the Titan X Pascal GPU used for this research.

REFERENCES

[1] Córcoles, J. I., Ortega, J. F., Hernández, D., and Moreno, M. A., "Estimation of leaf area index in onion (allium cepa l.) using an unmanned aerial vehicle," *Biosystems engineering* **115**(1), 31–42 (2013).

[2] Lazicki, P., Geisseler, D., and Horwath, W. R., "Onion Production in California." CDFA, June 2016 https://apps1.cdfa.ca.gov/FertilizerResearch/docs/Onion_Production_CA.pdf. (Accessed: 16 August 2018).

[3] Drinkwater, W. and Janes, B., "Effects of irrigation and soil moisture on maturity, yield and storage of two onion hybrids," in [*Proceedings of the American Society for Horticultural Science*], **66**, 267–278 (1955).

[4] Doneen, L. D. and MacGillivray, J. H., [*Suggestions on irrigating commercial truck crops*], University of California, College of Agriculture, Agricultural Experiment Station (1943).

[5] Singh, R. and Alderfer, R., "Effects of soil-moisture stress at different periods of growth of some vegetable crops," *Soil Science* **101**(1), 69–80 (1966).

[6] Govender, M., Govender, P., Weiersbye, I., Witkowski, E., and Ahmed, F., "Review of commonly used remote sensing and ground-based technologies to measure plant water stress," *Water SA* **35**(5) (2009).

[7] Zhao, T., Stark, B., Chen, Y., Ray, A. L., and Doll, D., "A detailed field study of direct correlations between ground truth crop water stress and normalized difference vegetation index (ndvi) from small unmanned aerial system (suas)," in [*Unmanned Aircraft Systems (ICUAS), 2015 International Conference on*], 520–525, IEEE (2015).

[8] Zhao, T., Stark, B., Chen, Y., Ray, A., and Doll, D., "More reliable crop water stress quantification using small unmanned aerial systems (suas)," *IFAC-PapersOnLine* **49**(16), 409–414 (2016).

[9] Zhao, T., Stark, B., Chen, Y., Ray, A. L., and Doll, D., "Challenges in water stress quantification using small unmanned aerial system (suas): Lessons from a growing season of almond," *Journal of Intelligent & Robotic Systems* **88**(2-4), 721–735 (2017).

[10] Zhao, T., Doll, D., Wang, D., and Chen, Y., "A new framework for uav-based remote sensing data processing and its application in almond water stress quantification," in [*Unmanned Aircraft Systems (ICUAS), 2017 International Conference on*], 1794–1799, IEEE (2017).

[11] Zhao, T., Chen, Y., Ray, A., and Doll, D., "Quantifying almond water stress using unmanned aerial vehicles (uavs): correlation of stem water potential and higher order moments of non-normalized canopy distribution," in [*ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*], V009T07A058–V009T07A058, American Society of Mechanical Engineers (2017).

[12] Zhao, T., Doll, D., and Chen, Y., "Better almond water stress monitoring using fractional-order moments of non-normalized difference vegetation index," in [*2017 ASABE Annual International Meeting*], 1, American Society of Agricultural and Biological Engineers (2017).

[13] Kim, Y., Glenn, D. M., Park, J., Ngugi, H. K., and Lehman, B. L., "Hyperspectral image analysis for water stress detection of apple trees," *Computers and Electronics in Agriculture* **77**(2), 155–160 (2011).

[14] Maimaitiyiming, M., Ghulam, A., Bozzolo, A., Wilkins, J. L., and Kwasniewski, M. T., "Early detection of plant physiological responses to different levels of water stress using reflectance spectroscopy," *Remote Sensing* **9**(7), 745 (2017).

[15] González-Fernández, A. B., Rodríguez-Pérez, J. R., Marcelo, V., and Valenciano, J. B., "Using field spectrometry and a plant probe accessory to determine leaf water content in commercial vineyards," *Agricultural water management* **156**, 43–50 (2015).

[16] Marino, S., Basso, B., Leone, A., and Alvino, A., "Agronomic traits and vegetation indices of two onion hybrids," *Scientia Horticulturae* **155**, 56–64 (2013).

[17] Ballesteros, R., Ortega, J. F., Hernandez, D., and Moreno, M. A., "Onion biomass monitoring using uav-based rgb imaging," *Precision Agriculture* , 1–18 (2018).

[18] Wang, W., Li, C., and Gitaitis, R. D., "Optical properties of healthy and diseased onion tissues in the visible and near-infrared spectral region," *Transactions of the ASABE* **57**(6), 1771–1782 (2014).

[19] Wang, W. and Li, C., "A multimodal machine vision system for quality inspection of onions," *Journal of Food Engineering* **166**, 291–301 (2015).

[20] Islam, M. N., Nielsen, G., Stærke, S., Kjær, A., Jørgensen, B., and Edelenbos, M., "Novel non-destructive quality assessment techniques of onion bulbs: a comparative study," *Journal of Food Science and Technology* , 1–11 (2018).

[21] "The worlds first material sensing smartphone."

[22] Jolliffe, I., "Principal component analysis," in [*International encyclopedia of statistical science*], 1094–1096, Springer (2011).

[23] Fisher, R. A., "The use of multiple measurements in taxonomic problems," *Annals of eugenics* **7**(2), 179–188 (1936).

[24] Rosenblatt, F., "The perceptron: a probabilistic model for information storage and organization in the brain.," *Psychological review* **65**(6), 386 (1958).

[25] Minsky, M. and Papert, S., "Perceptrons.," (1969).

[26] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al., "Scikit-learn: Machine learning in python," *Journal of machine learning research* **12**(Oct), 2825–2830 (2011).