AFRL-AFOSR-JP-TR-2019-0046

Autonomous Learning in a Dynamic World

Geoff Webb
MONASH UNIVERSITY
WELLINGTON RD
CLAYTON, 3800
AU

07/24/2019
Final Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
Air Force Office of Scientific Research

Asian Office of Aerospace Research and Development
Unit 45002, APO AP 96338-5002

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Executive Services, Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
24-07-2019	Final	14 Jul 2017 to 13 Jul 2019

4. TITLE AND SUBTITLE
Autonomous Learning in a Dynamic World

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA2386-17-1-4033

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Geoff Webb

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
MONASH UNIVERSITY
WELLINGTON RD
CLAYTON, 3800 AU

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AOARD
UNIT 45002
APO AP 96338-5002

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/AFOSR IOA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-AFOSR-JP-TR-2019-0046

12. DISTRIBUTION/AVAILABILITY STATEMENT
A DISTRIBUTION UNLIMITED: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Autonomous Learning in a Dynamic World developed algorithms to autonomously adapt to concept drift that is sufficiently robust for real-world applications. The authors focused finding the 'sweet path' across a 3-dim space of forgetting-rate, bias-variance, and drift-rate. A critical finding is that drift may be domain or application dependent--thus, concept drift mapping is a requisite for adaptations. This was shown through an airlines dataset. An appropriate mapping method is then possible after characterizing the drift in the data. Using these approaches, the authors implemented this understanding to develop the novel Hoeffding Anytime Tree as and advancement over the state-of-the-art. The Hoeffding Anytime Tree splits a leaf when it has statistical evidence that the best potential split is better than no split and continually monitors the performance of the split and potential alternatives and replaces the split when it has statistical evidence that the alternative is better. This was then tested using the UCI human activity recognition data set, showing up to an order of magnitude decrease in error rate. This project resulted in 3 peer-reviewed journals, 1 conference paper, and 1 arxiv paper.

15. SUBJECT TERMS
Autonomous learning, Concept drift, Attribute dependency, Bayesian classifier, Logistic regression, Bias-variance tradeoff, Classifier ensemble

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			LIN, ALAN
Unclassified	Unclassified	Unclassified	SAR		19b. TELEPHONE NUMBER *(Include area code)* 315-227-7009

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Final Report for AOARD Grant FA2386-17-1-4033

"Autonomous Learning in a Dynamic World"

15 July 2019

Name of Principal Investigators (PI and Co-PIs): Geoffrey Ian Webb
- e-mail address : geoff.webb@monash.edu
- Institution : Monash University
- Mailing Address : Wellington Road, Clayton, VIC 3800, Australia
- Phone : +61 3 99053296
- Fax : +61 3 99055159

Period of Performance: 07/26/2017 - 07/25/2019

Abstract:

The world is dynamic, in a constant state of flux, but machine learning typically learns static models from historical data. As the world changes, these models can lose veracity, declining in utility, sometimes precipitously so. This project develops new machine learning technologies to ameliorate this problem. We have made progress on all of theory, methodology and techniques. In terms of theory, we have posited the existence of a critical link between drift rate, ideal forgetting rate, and ideal bias-variance learning profile. In terms of methodology, we have developed practical methods for analyzing real-world drift from sample data. In terms of techniques, we have developed a new incremental decision tree learning algorithm that is statistically more powerful than the previous state-of-the-art, and much more responsive to drift.

Introduction:

The world is dynamic, and data about how the world used to be are not always pertinent to how the world currently is. For example, if traffic laws change, models that were appropriate given the old laws may not work with the new. As another example, as the culture of road behavior changes, models adapted to old norms may not be optimal under the new. As per the first example, some changes may be abrupt and, as per the second, some may be gradual. As per the first, some changes may be predictable and, as per the second, some may be unpredictable.

There has been much research into learning in the context of concept drift (as such change is called in machine learning) [1-4]. However, these academic exercises are not yet of sufficient robustness for real world deployment. This is attested to by the fact that, in practice, the issue of drift is typically addressed by having a regular schedule (e.g. weekly) on which the oldest data is discarded from the reference data used for learning (the training set), the most recent data added and a new model learned.

However, this approach is suboptimal. While drift is usually conceived of as a uniform process, there is good reason to believe that different parts of the data space will drift in substantially different ways. For example, a change in the law about minimum distances that must be maintained between vehicles will change the distance a vehicle should expect to be from other vehicles, with consequent implications for maintenance of safe vehicle speed. However, it will not change the relevance of evidence obtained from sensors about road condition. Hence, while the aspect of the speed maintenance system relating to vehicle distance should be relearned, the long history of data relating to road condition should all be

DISTRIBUTION A Distribution Approved for Public Release: Distribution Unlimited

retained and utilized to learn the best possible model.

This project addresses autonomous machine learning in the context of concept drift. It is developing new theory, methods and algorithms that exploit our recent insights and advances in learning technology to allow learning algorithms to adapt dynamically and appropriately to complex change without human intervention.

AIMS: This project will investigate the hypothesis that different forms of concept drift are best addressed by different combinations of data treatments and learner bias-variance profiles. It will also investigate the further hypothesis that concept drift often varied in rate and form in different data subspaces and that different pairs of data treatment and learner bias-variance profile will thus be most effective in different data subspaces at any given point in time. It will develop algorithms informed by these hypotheses and assess their performance across a wide range of forms of concept drift. Our ultimate aim is to develop the first incremental learning algorithms that autonomously adapt to concept drift in different ways in the context of different forms of drift in different data subspaces.

Method/Theory/Experiment/Results and Discussion:

We pursue a three-pronged approach, working first of theory and then using our theory to drive advances in both methodology and techniques.

Theory

We posit relationships between types of concept drift and the properties of the learners that best handle those forms of drift. Specifically, we propose and investigate two hypotheses –

> *The drift-rate/forgetting-rate nexus*. As the rate of concept drift increases, model accuracy will in general be maximized by increasing forgetting rates, and conversely, as the drift rate decreases, model accuracy will in general be maximized by decreasing forgetting rates. Here *increasing forgetting rates* means focusing on more recent evidence by reducing window sizes or increasing decay rates. *Decreasing forgetting rates* means focusing on longer term evidence by increasing window sizes or decreasing decay rates.
>
> *The forgetting-rate/bias-variance-profile nexus*. As forgetting rates increase, model accuracy will in general be maximized by altering the bias/variance profile of a learner to decrease variance, and conversely, as forgetting rates decrease, model accuracy will in general be maximized by decreasing bias.

The first of these hypotheses is intuitive. The faster the world changes; the less relevance older information will have. In consequence, more aggressive forgetting mechanisms, specifically smaller windows or higher decay rates, will be required to exclude older examples for which the trade-off between providing additional information that is relevant to the current state-of-the-world and that which is misleading is weighted too heavily to the latter.

The second hypothesis derives from another hypothesis; that when learning from smaller quantities of data, lower variance learners will maximize accuracy due to their ability to avoid overfitting, whereas when learning from larger datasets lower bias learners will maximize accuracy, due to their ability to model the details present in large data [5]. The more past data we forget, the smaller the effective data quantity from which we learn.

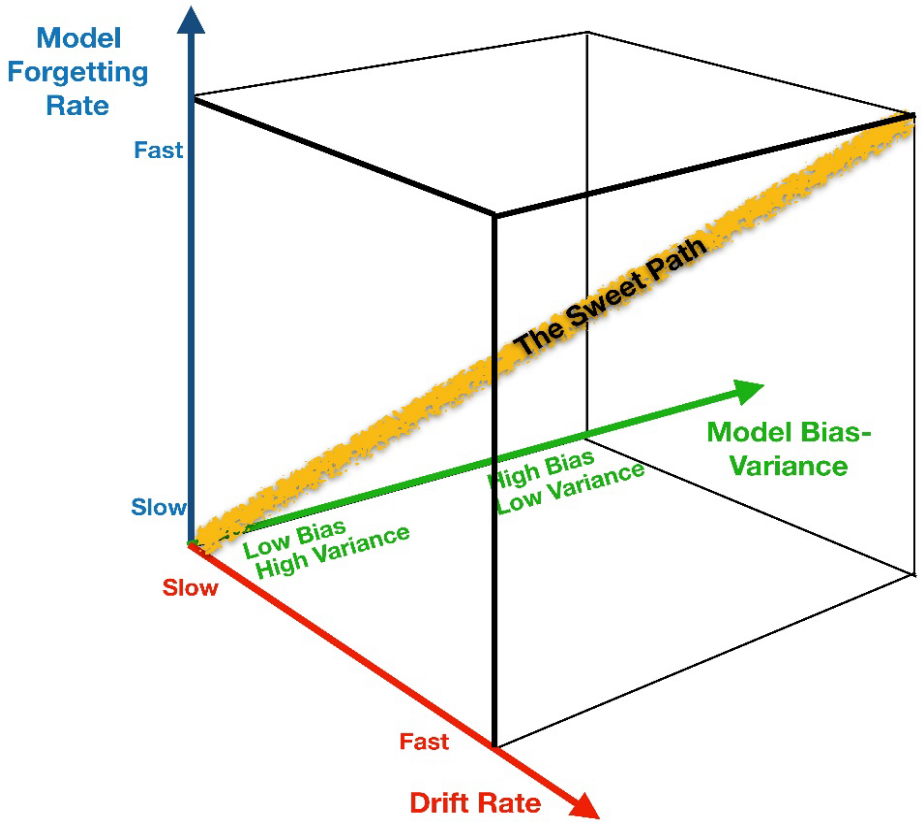

Figure 1: The Sweet Path. For a data stream with slow drift, error will be minimised by learning with a low bias learner with a slow forgetting rate. As drift rate increases, so too will the optimal forgetting rate and the ideal bias/variance profile will more greatly favor lower variance.

Therefore, with high-forgetting rate, low-variance models are more desirable and low-bias models with low-forgetting rate.

In conjunction, these hypothesized effects imply the *sweet path* for concept drift illustrated in Figure 1. On this sweet path, the lowest error for a low drift rate will be achieved by a low bias learner with a low forgetting rate and the lowest error for a high drift rate will be achieved by a low variance learner with a high forgetting rate.

To explore the drift-rate/forgetting-rate/bias-variance-profile nexus, we wish to systematically manipulate the rate of concept drift in sample data and observe whether the lowest predictive error with fast drift results from fast forgetting coupled with low variance while the lowest predictive error with slow drift results from slow forgetting coupled with low bias. To this end, we require an incremental learner that can learn from sliding windows or with decay. We also require means of varying the learner's bias/variance profile.

For our experiments, we use the semi-naive Bayesian method AnDE [6], as it satisfies these requirements. First, the model has a tuneable parameter n that controls the representation bias and variance. When n=0 (in AnDE), one gets a naive Bayes classifier which is highly biased but has low variance. Higher values of n decrease bias, at the cost of an increase in variance. Lower values of n decrease variance, at a cost of increased bias. Second, the AnDE model can be represented using counts of observed marginal frequencies, the dimensionality of each of which is controlled by n. These can readily be incrementally updated to reflect a sliding window or incremental decay without need for relearning the entire model.

We use Δ=0.05 to represent fast drift, Δ=0.01 for medium drift and Δ=0.0005 for slow drift.

We selected these from a wider range of values explored, as exemplars that demonstrate clear differences in outcomes.

We present here results for manipulating forgetting rate using decay. The results (presented in the paper) when windows are used for forgetting are consistent with the results for decay.

We generate data streams of 5,000 successive time steps, at each time step drawing one example randomly from the probability distribution for that step and drifting the distribution every 10 steps.

We use prequential evaluation, whereby at each time step the current model is applied to classify the next example in the data stream and then the example is used to update the model. We run each experiment 150 times for NB and A1DE and 100 times for A2DE (due to there being insufficient time to complete more runs). We present averages over all runs. Figures 1 to 9 plot the resulting error rates, where each point in the plot is the average error over 50 successive time steps. Each plot presents the prequential error over successive time steps for a single classifier and single drift rate with each of slow (0.005), medium (0.05) and fast (0.15) decay. Figure 1 shows naïve Bayes with slow drift through to Figure 9 which shows A2DE with fast drift.

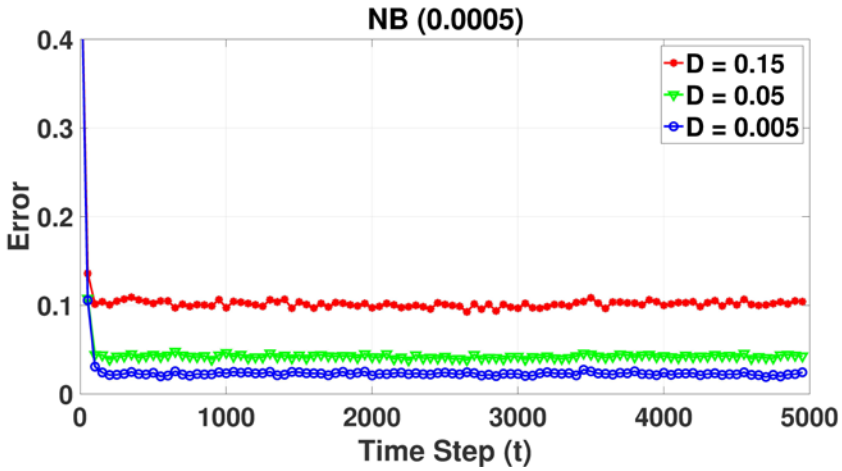

Figure 1 Naive Bayes (low variance) with slow drift

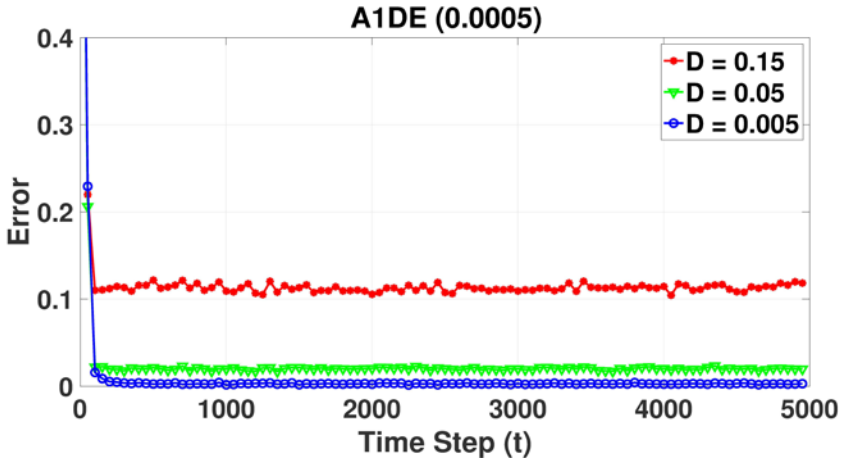

Figure 2 A1DE (intermediate bias/variance) with slow drift

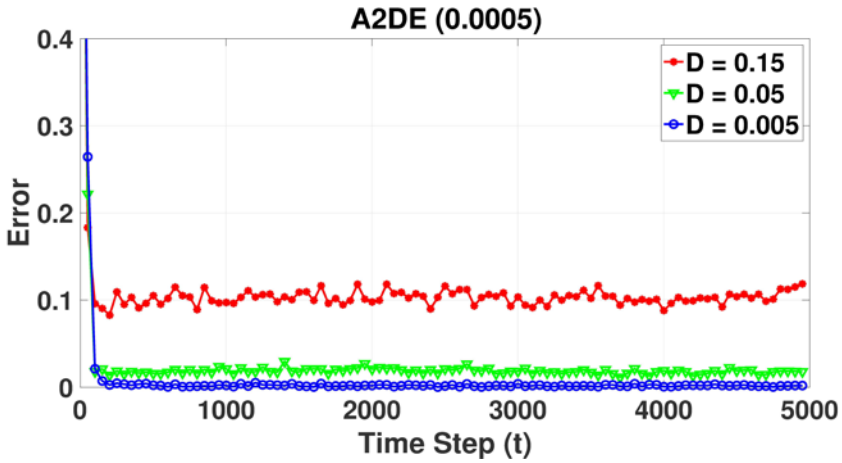

Figure 3 A2DE (low bias) with slow drift

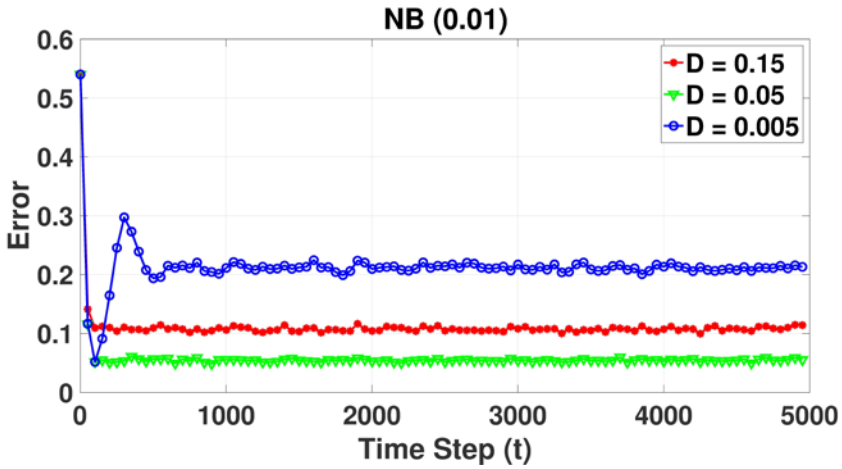

Figure 4 Naive Bayes (low variance) with intermediate drift

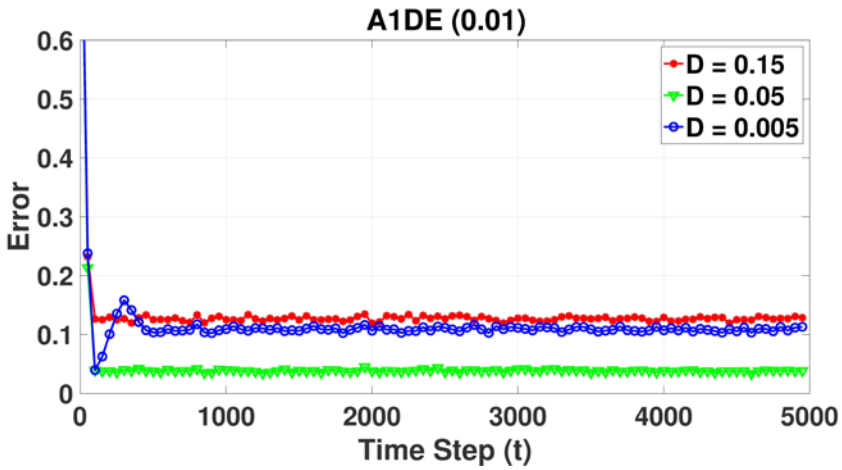

Figure 5 A1DE (intermediate bias/variance) with intermediate drift

DISTRIBUTION A Distribution Approved for Public Release: Distribution Unlimited

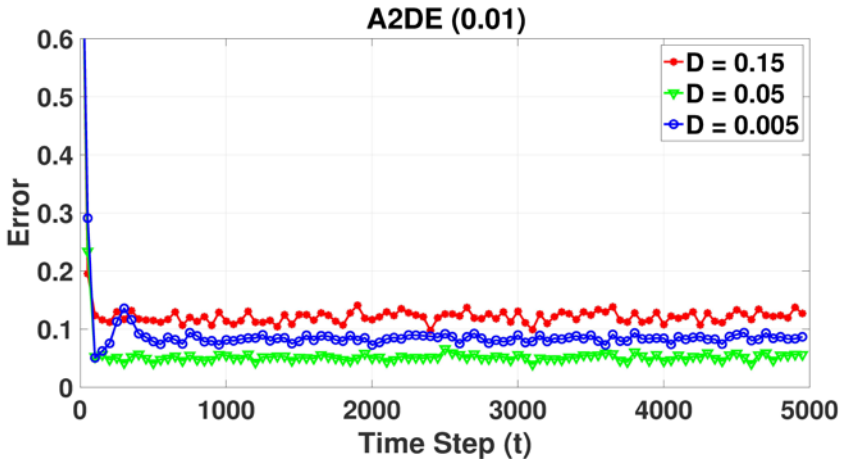

Figure 6 A2DE (low bias) with intermediate drift

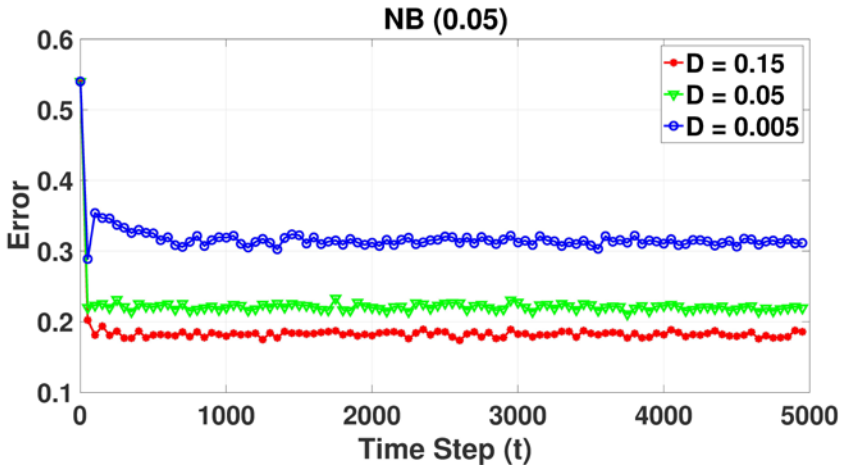

Figure 7 Naive Bayes (low variance) with fast drift

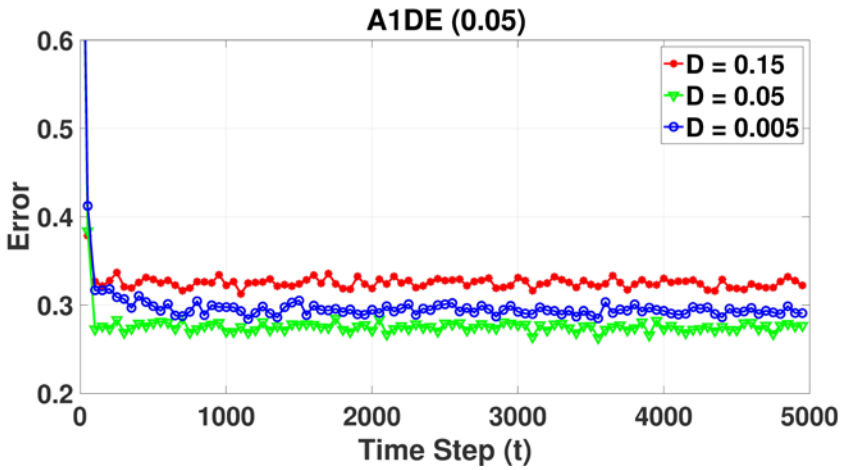

Figure 8 A1DE (intermediate bias/variance) with fast drift

DISTRIBUTION A Distribution Approved for Public Release: Distribution Unlimited

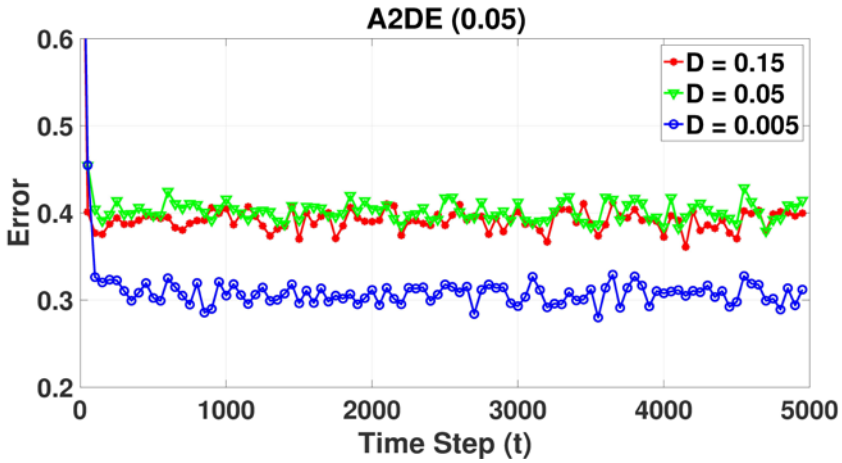

Figure 9 A2DE (low bias) with fast drift

For slow drift (Δ=0.0005), all three learners obtain the lowest error with the slowest forgetting rate (smallest decay rate). The lowest bias learner (A2DE) achieves the lowest error. This is consistent with our two hypotheses, and establishes the slow drift end of the sweet path.

At an intermediate drift rate (Δ = 0.01), the intermediate bias learner A1DE outperforms NB. For this intermediate drift rate the intermediate decay rate (0.05) attains the lowest error and the learner with intermediate bias (A1DE) minimizes overall error. These results for the intermediate drift rate are also consistent with our hypotheses. For intermediate drift the lowest error is obtained with an intermediate forgetting rate and an intermediate bias/variance trade-off. This establishes the middle section of the sweet path.

For fast drift NB achieves its lowest error fast forgetting (decay rate 0.15).

However, contrary to our expectations, A1DE and A2DE achieve their lowest error with slower forgetting. This is because these models effectively fail in the face of such rapid drift. For technical reasons explained in the paper, the relative error is dominated by the ability to accurately estimate three low order invariant probabilities and the performance approximates learning from a stationary distribution when the majority of attributes are noise attributes.

Nonetheless, the lowest overall error is obtained by the fastest forgetting rates and the lowest variance learner (NB). Thus, while the results for A1DE and A2DE may initially appear to be an anomaly, the results for fast drift are also consistent with our hypotheses and establish the fast drift end of the sweet path.

These results are highly significant, as they imply that an incremental learner that may confront varying rates of concept drift ought to be able to change one of its fundamental properties, its bias-variance profile.

For more details please consult our paper “On the Inter-Relationships among Drift Rate, Forgetting Rate, Bias/Variance Profile and Error.” [P5]

Methodology

In order to manage concept drift, it will often be valuable to understand the exact nature of the drift that affects the relevant application. To this end, we propose a new data mining task, *concept drift mapping* – the description and analysis of instances of concept drift. We argue that concept drift mapping is an essential prerequisite for tackling concept drift. We propose tools for this purpose, arguing for the importance of quantitative descriptions of

drift and shift in marginal distributions. We present quantitative concept drift mapping techniques for categorical data, along with methods for visualizing their results. We illustrate their effectiveness with real-world case studies across energy-pricing, vegetation monitoring and airline scheduling.

The airlines case study provides a good representation of the insights the techniques can provide. Each example in this data represents a flight, with covariates *Airline*, *Flight*. *AirportFrom*, *AirportTo*, *DayOfWeek*, *Time*, and Length and with a binary class indicating whether the flight arrived on time. *DayOfWeek* has been used to partition the data into days and weeks and has not been included as a covariate in the analysis. Figure 10 shows the covariate drift from day to day. Figure 11 shows the covariate drift for the week prior to a day against the week starting with that day and is plotted daily from the seventh day. Note that the numbering starts with 4 as the first day in the data is day number 3.

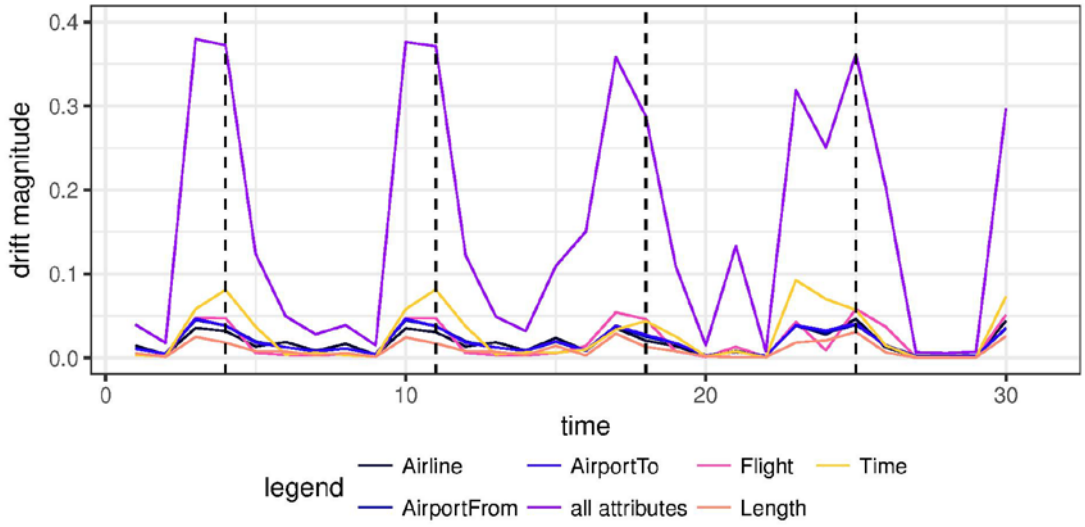

Figure 10 Daily covariate drift

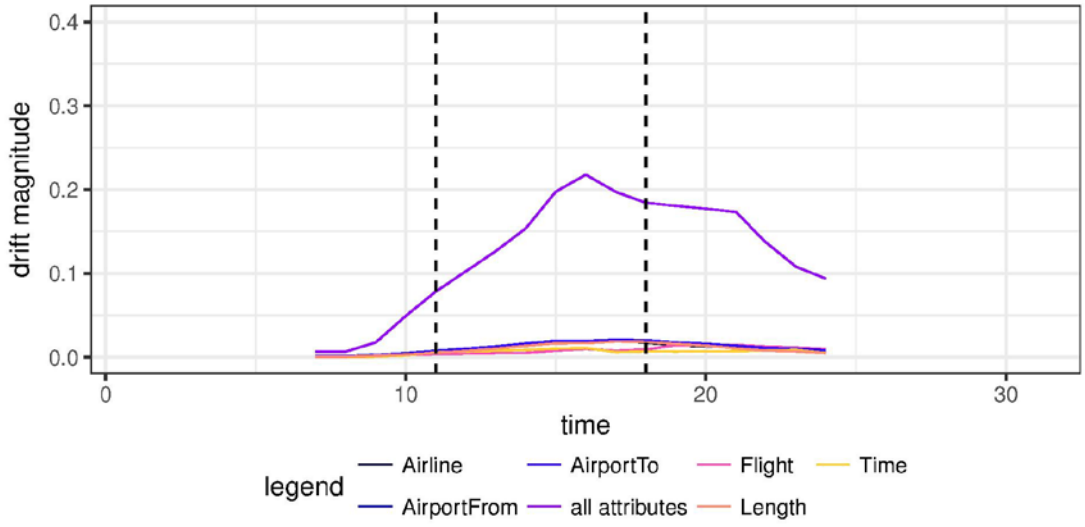

Figure 11 Weekly covariate drift

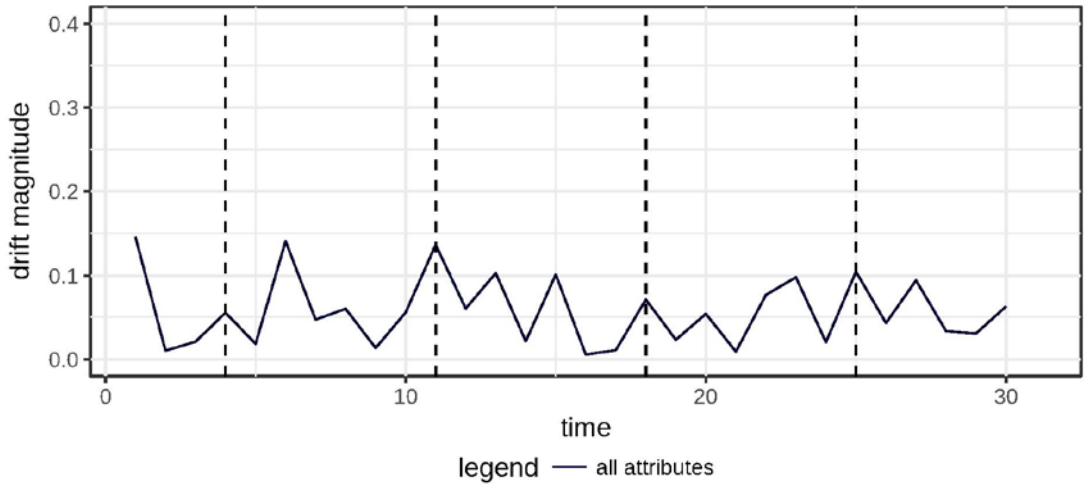

Figure 12 Daily class drift

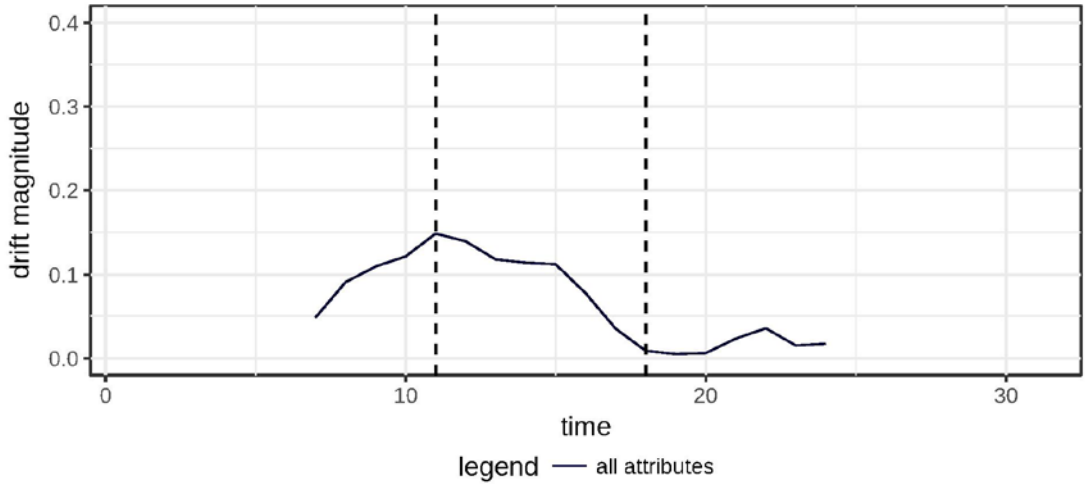

Figure 13 Weekly class drift

The first figure shows that for the first two weeks there is a cyclical pattern in the magnitude of covariate drift, with large changes from Friday to Saturday and from Saturday to Sunday, but lower drift from Sunday to Monday and substantially lower drift between successive weekdays. However, this pattern breaks down over the following two weeks. Unfortunately we do not have the dates for which the data were collected and hence can only speculate for the reasons for this change in pattern; weather and public holidays being two potential explanations. The marginal distributions indicate that the time of day is the major contributor to drift for most of the period but that flight number overtakes it for some parts of the second half of the period.

The weekly analysis shows that while there is substantial drift from day to day, there is little drift between the first two weeks, confirming the notion that they follow a steady cycle. The inter-week drift then rises sharply. Interestingly, it is the origin and destination airports and flight lengths that change most from week to week as opposed to the time of day and flight number which dominated the inter-day drift.

Figure 12 and Figure 13 show the daily and weekly class drift, respectively. They reveal that the class, representing on-time performance, is not subject to the same weekly cycle of drift as the covariates and that there is greatest drift in on-time performance between the second and third weeks. It is interesting to contrast the inter-week covariate drift to the

inter-week class drift. The covariates start with almost no drift which then increases substantially, while the class starts with substantial drift and subsequently drops to having almost no drift. In general, these plots are revealing in that they show that the class drift for this data is quite different in nature to the covariate drift.

This data demonstrates the importance of the granularity of the time periods used in drift analysis and the manner in which different granularities can each convey different and valuable insights. It also illustrates how it is revealing to consider each of the different forms of drift, joint, class, covariate, conditioned class and conditioned covariate. These different aspects of a distribution may each drift in different ways, and an analysis that does not consider all may miss important insights into the nature of drift in a domain.

A detailed description of this research is provided in the paper, "Analyzing concept drift and shift from sample data." [P3]

We also investigate how drift mapping might be extended to numeric data. To this end we assess the applicability of each of the various methods that have been proposed for measuring distance between numeric distributions. Our findings are as follows.

1. While Hellinger distance, Kullback–Leibler divergence, Kolmogorov–Smirnov distance and total variation distance can be numerically approximated for univariate data, they do not scale to higher dimensions.
2. Kolmogorov–Smirnov distance cannot be applied to data with more than 2 variables.
3. The exponential complexity of Hellinger distance, Kullback–Leibler divergence and total variation distance limits their scalability.
4. Hellinger distance and Kullback–Leibler divergence have closed-form solutions for some known distributions, including the multivariate normal. In consequence, these measures can be used for multidimensional data when a normal distribution can be assumed.
5. Hellinger distance and total variation distance have three advantages relative to Kullback–Leibler divergence:
    a. They are metrics, while Kullback–Leibler divergence is not.
    b. Their numerical approximation is more robust and less sensitive to selected parameters.
    c. They return a value, bounded between 0 and 1, that is commensurable from application to application, while the Kullback–Leibler divergence is unbounded and incommensurable between applications.

Our recommendations are as follows.

1. For univariate or low-dimensional numeric data Hellinger distance is an effective measure of dissimilarity between distributions.
2. T-statistics provide an effective unitless approximation of the distance between sample means for univariate numeric data.
3. Hotelling T^2 provides an effective unitless approximation of the distance between sample means for multivariate numeric data.
4. Unless it is possible to assume the data are drawn from a known distribution, such as the normal distribution, there are few effective means of measuring distance between samples of high-dimensional numerical data. The PCA-based approach provides the most effective current method, but warrants further investigation.

A detailed description of this research is provided in our paper "Survey of distance measures for quantifying concept drift and shift in numeric data." [P1]

Techniques

Deep learning has proved extremely effective at learning in a wide variety of contexts where the relevant features are not well understood. There has been little research into methods for managing deep learning in the context of concept drift. To this end we explore how concept drift maps might be deployed to adapt deep learning as drift occurs. Specifically, we investigate an adaptive extreme learning machine that uses a concept drift map to regulate the forgetting factor. It does this by first estimating the distribution of each class on each attribute each time a batch of new instances are labeled. This is used to estimate the magnitude of concept drift relative to the distribution on the last batch. This drift estimate is then used to regulate the forgetting factor, which in turn regulates updates to the model. Experiments on benchmark stream learning data sets demonstrate the proposed model is generally more accurate than previous approaches in the presence of drift.

A detailed description of this research is provided in the paper, "Adaptive Online Extreme Learning Machine by Regulating Forgetting Factor by Concept Drift Map." [P2]

The Hoeffding Tree is an incremental decision tree learning algorithm that has become the workhorse of the machine learning with concept drift community. We introduce a novel incremental decision tree learning algorithm, Hoeffding Anytime Tree, that is statistically more efficient than the current state-of-the-art, Hoeffding Tree. Hoeffding Anytime Tree differs from Hoeffding Tree only in two respects. The first is that Hoeffding Tree converts a leaf into a decision tree split when it has statistical evidence that the best potential split is better than the second best. In contrast, Hoeffding Anytime Tree splits a leaf when it has statistical evidence that the best potential split is better than no split.

The second difference is that once a leaf has been converted to a split, Hoeffding Tree does not consider replacing the split, whereas Hoeffding Anytime Tree continually monitors the performance of the split and potential alternatives and replaces the split when it has statistical evidence that the alternative is better.

We demonstrate that an implementation of Hoeffding Anytime Tree – "Extremely Fast Decision Tree," a minor modification to the MOA implementation of Hoeffding Tree – obtains significantly superior prequential accuracy on most of the largest classification datasets from the UCI repository. Figure 14 shows their performance on the largest UCI classification dataset, human activity recognition, when the dataset is shuffled (and hence there is no concept drift). Notably, even after learning from 30 million examples, EFDT is orders of magnitude more accurate than VFDT. Figure 15 shows performance when the dataset is not shuffled. EFDT responds much more effectively to the sudden drift imposed by the dataset's native order, with an overall error rate of less than 0.2% compared to VFDT's error rate of more than 4.0%. These results are typical of those for the large natural datasets in the UCI repository.

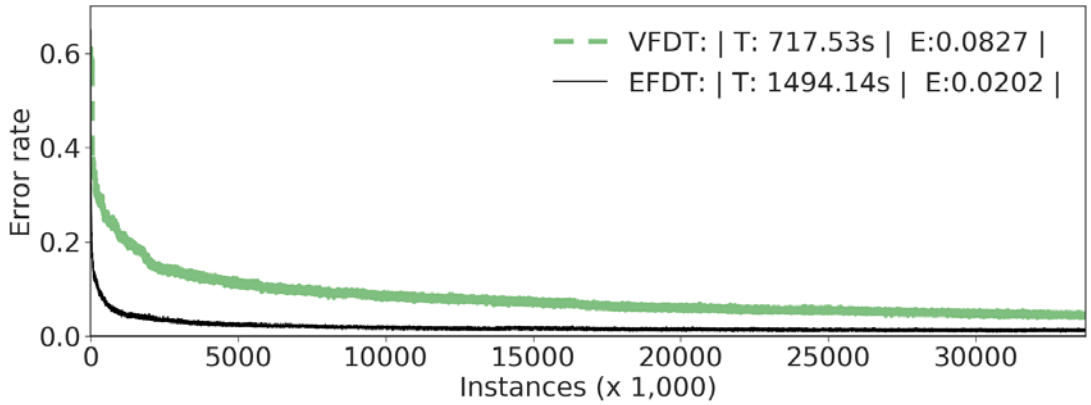

Figure 14 Average prequential error of EFDT and VFDT on shuffled human activity recognition dataset.

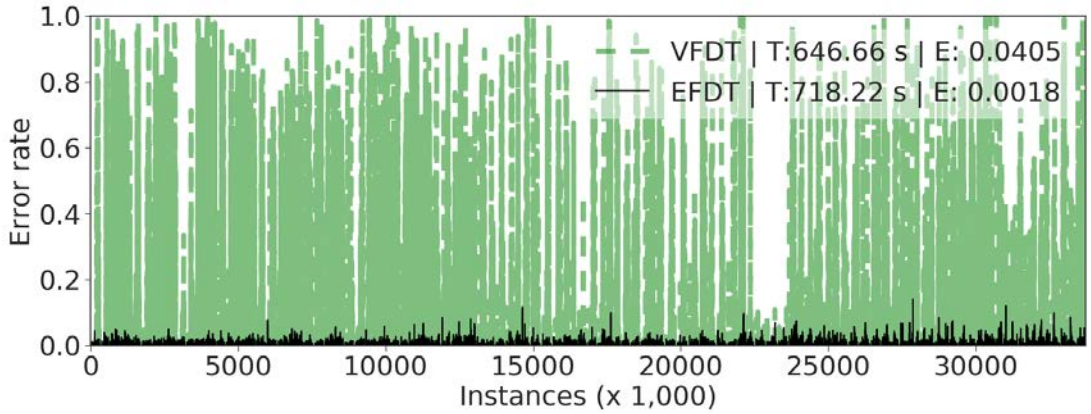

Figure 15 Prequential error of EFDT and VFDT on the unshuffled human activity recognition dataset.

Hoeffding Anytime Tree produces the asymptotic batch tree in the limit, is naturally resilient to concept drift, and can be used as a higher accuracy replacement for Hoeffding Tree in most scenarios, at a small additional computational cost. Our preliminary results indicate that it is especially effective in the face of concept drift due to its capacity to replace subtrees that are no longer relevant.

A detailed description of this research is provided in the paper, “Extremely Fast Decision Tree.” [P4]

Conclusions

We have investigated three facets of the problems confronting machine learning in a dynamic and ever-changing world. We have developed the *sweet path* theory, that there is in inextricable connection between the trio of drift rate, forgetting rate and ideal bias-variance trade-off. This has the

References

[1] Gama, J., Zliobaite, I., Bifet, A., Pechenizkiy, M. and Bouchachia, A. A Survey on Concept Drift Adaptation. *Acm Comput Surv*, 46, 4 (Mar 2014).

[2] Gao, J., Fan, W., Han, J. and Philip, S. Y. *A General Framework for Mining Concept-Drifting Data Streams with Skewed Distributions*. SIAM, City, 2007.

[3] Sammut, C. and Harries, M. *Concept Drift*. Springer, City, 2010.

[4] Webb, G. I., Hyde, R., Cao, H., Nguyen, H. L. and Petitjean, F. Characterizing Concept Drift. *Data Mining and Knowledge Discovery*, 30, 4 (2016), 964-994.

[5] Brain, D. and Webb, G. I. *The need for low bias algorithms in classification learning from large data sets*. Springer-Verlag, City, 2002.

[6] Webb, G. I., Boughton, J., Zheng, F., Ting, K. and Salem, H. Learning by extrapolation from marginal to full-multivariate probability distributions: decreasingly naive Bayesian classification. *Machine Learning*, 86, 2 (2012), 233-272.

List of Publications and Significant Collaborations that resulted from your AOARD supported project:

a) papers published in peer-reviewed journals,

[P1] **Survey of distance measures for quantifying concept drift and shift in numeric data.**
Goldenberg, I., & Webb, G. I.
Knowledge and Information Systems, in press.

[P2] **Adaptive Online Extreme Learning Machine by Regulating Forgetting Factor by Concept Drift Map.**
Yu, H., & Webb, G. I.
Neurocomputing, 343, 141-153, 2019.

[P3] **Analyzing concept drift and shift from sample data.**
Webb, G. I., Lee, L. K., Goethals, B., & Petitjean, F.
Data Mining and Knowledge Discovery, 32(5), 1179-1199, 2018.

b) papers published in peer-reviewed conference proceedings,

[P4] **Extremely Fast Decision Tree.**
Manapragada, C., Webb, G. I., & Salehi, M.
Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, pp. 1953–1962, 2018.

c) papers published in non-peer-reviewed journals and conference proceedings,
d) conference presentations without papers,
e) manuscripts submitted but not yet published

[P5] **On the Inter-Relationships among Drift Rate, Forgetting Rate, Bias/Variance Profile and Error.**
Zaidi, N. A., Webb, G. I., Petitjean, F., & Forestier, G.
arxiv, 1801.09354, 2018.

f) provide a list any interactions with industry or with Air Force Research Laboratory scientists or significant collaborations that resulted from this work.

Attachments: Publications a), b) and c) listed above if possible.

DD882, SF425: As a separate document, please complete the invention disclosure form

(**DD882)**, Federal financial report (**SF425**) and sign both. **SF425 must be signed by an authorized person in your business office.**

Important Note: **Attached publications are used only for internal use.** They do not go outside of AFRL because of the copyright issues of the published papers. However, **the main text goes to DTIC which can be accessible to public.** Thus, **a final report must be self-contained without reference to other documents. Submission of a report that is very similar to a full length journal article will be sufficient in most cases. The final report should give a fair account of the work performed during the period of performance.** There will be variations depending on the scope of the work. As such, there is no length or formatting constraints for the final report. Keep in mind the amount of funding you received relative to the amount of effort you put into the report. For example, do not submit a $300k report for $50k worth of funding; likewise, do not submit a $50k report for $300k worth of funding. **Include as many charts and figures as required to explain the work.**