P1.1 TEACHING ARTIFICIAL INTELLIGENCE TO METEOROLOGY UNDERGRADUATES

George S. Young* and Sue Ellen Haupt

The Pennsylvania State University, University Park, Pennsylvania

1. INTRODUCTION

Advanced Artificial Intelligence (AI) forecast systems have shown superiority to traditional regression models in a number of research projects (e.g. Burrows 1997, Dean and Fiedler 2002). Yet operational use of AI-based forecast systems remains limited in comparison with older regression-based statistical forecast systems. Likewise, development of AI forecast systems by operational forecasters remains uncommon. Given the potential of AI in operational forecasting why has adoption not been more widespread?

A major factor restricting the adoption of AI systems in operational meteorology is the lack of a pool of forecasters trained in the areas required to use the new systems wisely. The supply of forecasters and researchers trained in the theory and tools required to develop new AI-based forecast systems is even more limited. In order to develop a large enough pool of appropriately trained operational meteorologists one must teach AI at the undergraduate level.

The goal of this project was to develop an undergraduate course that would prepare students to be both users and developers of AI-based forecast systems. Successful users of such systems must understand the strengths and limitations of the various AI methodologies and understand verification well enough to be able to test for themselves the skill of AI systems versus their own forecasts. System developers require even greater knowledge, but not nearly as much as those researchers developing new AI methodologies.

As the development tools improve, system development is opened up to a much wider group of meteorologists. To use these tools wisely, the students must know enough theory to select the set of methods that are likely to work well on a particular problem. Likewise, to fit system development into an already busy schedule, operational meteorologists require both training and hands-on experience in a tool set that lets them develop AI forecast systems efficiently and within the limits of their knowledge.

Teaching AI to undergraduate meteorologists poses a number of challenges. In particular, undergraduate meteorologists are not fluent with all of the tools used by AI researchers. In mathematics they have generally had calculus and differential equations but often lack linear algebra. In statistics they understand quantitative data description and linear regression but have little theoretical background. Their computer skills are highly varied, often including programming ability in FORTRAN, C, or MATLAB, and virtually always including high level of familiarity with Excel. They generally lack knowledge of Prolog and other logic languages. Today's undergraduates are, however, more comfortable using GUI applications than in writing their own code, especially before taking the Department's required Computer Applications course.

These constraints dictate the instructional methods used to prepare our students. The text must cover the AI theory required by users and system developers, but should avoid the depth and complexity required by method developers. The hands-on part of the course should revolve around a GUI tool that lets students experiment with a broad range of AI methods to develop their own forecast systems. The tool should take full advantage of, and make high demands on, the students' knowledge of AI theory to achieve optimal system designs. The lecture component of the course should cover AI theory to just the level required to make wise and efficient use of the development tool and the AI-based forecast systems it produces. The laboratory and homework projects should exercise the students' knowledge of theory and show them that application of that knowledge results in more successful forecast systems. This report documents one approach to implementing these requirements in an undergraduate meteorology course.

2. RESOURCES

Teaching artificial intelligence to undergraduate meteorologists has become increasingly practical because of the recent explosion in the number and quality of AI resources aimed at applied scientists and business managers. The efforts of authors and development tool creators to reach the business market in particular have resulted in a wide choice of resources on a level appropriate to undergraduates. No longer are the available AI resources aimed predominantly at the researcher in computer science, statistics, or psychology. Instead the focus has shifted to a new generation of enabled users capable of developing and profiting from their own AI-based forecast systems.

2.1 *Texts*

The vast number of texts available in the AI arena can be divided loosely into two groups. Those of the first group are essentially research reports focusing on the authors' contributions to a particular sub-discipline of

* Corresponding author address: George S. Young, 503 Walker Building, Meteorology Department, The Pennsylvania State University, University Park, PA, 16802; e-mail: young@ems.psu.edu

AI. These texts are much more useful at the graduate level than as undergraduate tests. The second group includes general overviews aimed at the semi-technical reader. These texts are intended to prepare the reader to be a safe and effective developer and user of AI-based forecast systems. It is from this group that an undergraduate text should be selected.

Texts from both groups form an essential part of a professor's course preparation reading. While the overviews provide a good introduction and include much practical advice, they lack the theoretical depth required for developing robust lectures and answering student questions. The research reports that provide this depth are, however, much more approachable after one has perused an overview text.

In our opinion, one of the best texts for an undergraduate course in AI-based weather forecasting is "*Data Mining: Practical machine learning tools and techniques with JAVA implementations*" by Witten and Frank (2000). It is aimed at semi-technical audience with a mathematics background equivalent to our sophomores. It uses basic calculus to explain theory and ties in well with the Weka AI-system development tool distributed by the authors. The coverage is fairly strong on verification but weak on neural networks.

A possible alternative is "*Data Mining: Concepts, Models, Methods, and Algorithms*" by Kantardzic (2003). It is aimed at a fully technical audience but is still accessible to a junior or senior-level undergraduate meteorologist. As with the Witten and Frank text, Kantardzic uses basic calculus to explain theory. It is, however, weaker on forecast verification and lacks associated development tools. The two texts offer different enough perspectives on the same general set of AI methods as to provide highly complementary reading.

Neither text delves sufficiently into neural networks to fully support an undergraduate course in meteorological AI. Thus, a supplement such as "*Neural Smithing*" by Reed and Marks (1999) is required for course development. This single-topic book is quite readable but is aimed at a highly technical audience. While too deep and highly focused for use as an undergraduate text, it makes excellent reading and provides a sound basis for the development of a multi-week lecture series on neural networks and their application to meteorology.

2.2 *Development Tools*

Among the many commercial and freeware AI-development systems available today, one stands out. Weka, the free software companion to the Witten and Frank text, offers exceptional breadth, depth, and portability. Information about Weka and the software download site can be found on the world wide web at http://www.cs.waikato.ac.nz/~ml/. Weka is written in Java so it works transparently across all common operating systems. It uses a graphical user interface so it is approachable by the undergraduates. Perhaps most appealing, the broad range of AI methods implemented in Weka allow one to teach whatever methods are of interest whether or not they are prominently featured in the Witten and Frank text. Moreover, the easy access to new methods allow students to experiment with self teaching – ideal preparation for future on-the-job learning.

Another strong point of Weka is the well organized interface for setting the development parameters of each AI method. This easy access to the control parameters encourages students to experiment and allows the professor to design laboratory and homework exercises that demonstrate the benefits of theoretical understanding in the development and tuning of AI-based forecast systems.

Weka offers an adequate choice of verification methods for testing forecast system robustness but lacks many of the specialized verification diagnostics used in meteorology. These tools are easily implemented in Excel or an equivalent spreadsheet, so this is not a major shortcoming.

3. SYLLABUS

The syllabus for an undergraduate meteorology course in AI-based forecast system development is a moving target, evolving with the state of the art. The syllabus described here reflects the field in the Fall semester of 2003. The lecture topics reflect those AI methods that were either in widespread use or ready for transition from research to operations.

The first portion of the course covered the overall structure of an AI-based forecasting system and the pitfalls to be avoided in developing such a system. Lecture topics included forecast system dataflow, forecast system development data requirements, verification methods, and data quality control techniques.

The remainder of the course covered four general approaches to AI-based forecasting: classification trees, neural networks, advanced regression methods, and rule-based classification. The first two, classification trees and neural networks, have been demonstrated repeatedly in a research setting but have yet to make widespread inroads into operational weather forecasting. The third, advanced regression methods, covers issues of non-linearity and robustness that while routinely addressed in the mining of business data are not currently considered in most operational meteorological forecast systems. This discussion includes the fitting of non-linear models to improve the fit and the use of error measures other than least squares to reduce the impact of data errors and outliers. The last method, rule-based classification, offers strong possibilities for categorical forecasting of highly non-linear processes (e.g. road ice formation or tornado genesis), but lags even further in meteorological research and operations.

A total of 15 weeks (45 lectures and 45 one-hour laboratory periods) were available to teach this material. This was sufficient time to cover both the basic theory and the practical application of each of these four AI methods discussed. The laboratories and their associated homework exercises were an essential part of the teaching approach, providing a vivid illustration of the importance of theoretical understanding in the development and tuning of AI-based forecast systems. The use of Weka for the labs and homework allowed the students to work with multiple methods over the course of the semester, focusing on the understanding of AI issues rather than on the coding of individual routines.

The course involved four in-depth laboratory-homework assignments. The first illustrated the sensitivity of forecast system rankings to the choice of verification statistic. The insights to be gained by using more advanced diagnostic measures such as the skill score decompositions of Murphy (e.g. 1986, 1996) were also demonstrated. The second and third assignments involved the application of Weka to large datasets to develop working forecast systems.

In the second assignment, neural networks were used to forecast the surface flux of CO_2 from bulk meteorological measurements. The verification results for a comparison of two student neural network configurations against those of a linear regression model are shown in Table 1. The laboratory assignment stepped the students through a number of experiments to test the sensitivity of the skill on independent data to such neural network parameters as learning rate, momentum, and number of nodes and layers.

Table 1. Sample results from laboratory comparing linear regression to the Weka default neural network and the neural network produced by using two hidden layers, one of 4 nodes and the other of 2.

Score	Linear Regression	Default Neural Network	Tuned Neural Network
Correlation Coefficient	0.7789	0.8320	0.8484
Mean Absolute Error	3.0595	3.2707	2.8944
Root Mean Square Error	4.3738	4.3789	3.9796
Relative Absolute Error	57.38 %	61.34 %	54.28 %
Root Relative Square Error	65.11 %	65.19 %	59.25 %

In the third assignment, classification trees were used to forecast the occurrence of Instrument Flight Rules visibility (i.e. visibility < 3 miles) from a variety of synoptic data. The verification results for a comparison of two of the parameter sensitivity runs against linear regression are shown in Table 2. The other runs tested the sensitivity of skill on independent data on such classification tree parameters as the confidence threshold, the minimum number of cases per leaf, the use of pruning, and the use of subtree raising. As with the neural network assignment, the students discovered that appropriate tuning of these parameters was essential for the development of a forecast system capable of beating regression.

Table 2. Sample results from laboratory comparing logistic regression to two classification trees, one produced by the Weka default configuration and the other with the confidence threshold tightened from 0.25 to 0.125. The confusion matrix breaks down the results into a cross tabulation of forecasts and observations. The first column is for forecasts of occurrence and the second for forecasts of non-occurrence. Likewise, the first row is for observations of occurrence and the second is for observations of non-occurrence. A perfect forecast scheme would yield a diagonal confusion matrix.

Score	Logistic Regression	Default Classification Tree	Tuned Classification Tree
Correctly Classified Instances	93.91 %	93.34 %	94.07 %
Incorrectly Classified Instances	6.09 %	6.66 %	5.93 %
Confusion Matrix:	163 46 29 994	163 46 36 987	156 53 20 1003

Both the CO_2 and IFR problems are highly non-linear and so highlighted the need for advanced AI techniques. Likewise, both datasets included many extraneous predictors and thus tested the students' knowledge of overfitting and its prevention.

The final assignment was an AI forecast contest. The students were free to use whatever AI method they wished, provided that they could explain how it worked and how each control parameter was used to optimize its performance. The forecast parameter was a binary variable, the occurrence of Instrument Flight Rule ceilings (i.e. cloud base below 3000 feet above ground level).

Midterm and final exams were given as a further means of testing student understanding of AI methods and their pitfalls. As in the lectures and labs, the exams focused on the practical application of theory.

4. RESULTS

The course was first taught in the Fall semester of 2003 with overall positive results. Despite the potentially daunting mathematical bent, eleven students enrolled and all completed the semester successfully. All students were able to pass the theory-based exams at the method-user and system-developer levels. Likewise they had no trouble mastering the use of Weka for forecast system development via hands on instruction during the laboratory periods. Aided by group discussions and leading questions from the instructor, they were also successful in the application of theory to the choice of parameter values and to the understanding of their verification results.

While the exam grades tended to parallel the student's overall grade point average, the laboratory / homework assignments revealed a different pattern. Those students who showed the most interest and independence learned more and did much better on the assignments. The final forecast contest particularly highlighted this outcome with an enthusiastic, but only marginally prepared student taking the lead by trying methods that were not covered in class and then searching the web in order to obtain enough background material to use them effectively.

5. REFERENCES

Burrows, W. R. 1997: CART regression models for predicting UV radiation at the ground in the presence of cloud and other environmental factors. *J. Appl. Meteor*., **36**, 531–544.

Dean, A. R., and Fiedler, B. H. 2002: Forecasting warm-season burnoff of low clouds at the San Francisco International Airport using linear regression and a neural network. *J. Appl. Meteor*., **41**, 629–639.

Kantardzic, M., 2003: Data Mining: Concepts, Models, Methods, and Algorithms. Wiley-Interscience, 345pp.

Murphy, A. H. 1986: A New Decomposition of the Brier Score: Formulation and Interpretation. *Mon. Wea. Rev*., **114**, 2671–2673.

Murphy, A. H. 1996: General decompositions of MSE-based skill scores: Measures of some basic aspects of forecast quality. *Mon. Wea. Rev.,* **124**, 2353–2369.

Reed, R. E., and R. J. Marks, II, 1999: Neural Smithing, MIT Press, Cambridge, Massachusetts. 346pp.

Witten, I. H., and E. Frank, 2000: Data Mining: Practical machine learning tools and techniques with JAVA implementations. Morgan Kaufmann Publisher, San Francisco, California, 371pp.