

Contribution to the
Handbook of Data Mining and Knowledge Discovery
edited by W. Klösgen and J. Żytkow
New York, NY: Oxford University Press
(to appear)

**H1.3.1: Using decision tree induction to minimize process delays
in the printing industry**

Bob Evans R R Donnelley and Sons
Doug Fisher Vanderbilt University

Abstract

Rotogravure printing involves rotating a chrome-plated, engraved copper cylinder in a bath of ink, scraping off the excess ink, and pressing a continuous supply of paper against the engraved cylinder with a rubber roller, thus transferring ink from the engraved image of the cylinder to the paper.

The printing process is subject to many types of delays, one of which is *cylinder banding*. During the course of printing, grooves may become engraved into a cylinder surface. These grooves cause streaks or *bands* to be printed on the paper, thus ruining the final product.

This article describes the application of decision tree induction to identify the conditions under which banding did and did not occur at the Gallatin, Tennessee plant of R R Donnelley and Sons. Once found, discovered rules were used to bias printing press parameters towards conditions identified as favorable. This approach has been primarily responsible for reducing bands from 538 in 1989 to 26 in 1998. This article describes the technical and social issues addressed in the banding application.

H1.3.1.1 Process delay analysis using classification

In 1989 only weak human expertise existed regarding cylinder banding. Engineers had good ideas about the attributes relevant to banding, but their knowledge of attribute interactions was not sufficient to reliably predict, and thus guard against, banding incidents. A data mining tool was used to identify conditions under which banding was likely. In our case, each datum was

a ‘snapshot’ of press and environmental conditions that were true when a band occurred or that were true when one was reasonably sure that a band was not imminent. This latter case was not as clearcut as the former. In particular, print orders vary a great deal. Perhaps jobs with small print orders did not experience cylinder bands, but would if the cylinders continued to run. Fortunately, we found that nearly 90% of the cylinder bands occurred well before 1,000,000 copies were printed. To assign a *class* to a job, we used the following labeling rule. If a cylinder banded at any point, then we took a ‘snapshot’ of the system and labeled it as **Banded**. If a print order was greater than 1,000,000 copies and it ran to completion without a band, a snapshot of the final system was taken and labeled as **notBanded**. Cylinders that ran to completion in less than one million copies without a band were excluded as indeterminate. [\[link to Section C3.1\]](#)

Approximately 30 attributes described ‘snapshot’ (i.e., tuple). Each attribute represents a characteristic of a particular printing unit (e.g., **ink temperature**), cylinder (e.g., **chrome solution ratio**), printing press (e.g., **press speed**), paper (e.g., **paper type**), or more general environmental factors (e.g., **humidity**).

We used top-down induction of decision trees (TDIDT) (Quinlan 1991) to uncover diagnostic rules for cylinder banding. [\[link to Section C5.1.3\]](#) The final product of induction is a decision tree like the one illustrated in Figure 1, which was constructed over a sample of 177 tuples, 98 of which were labeled **notBanded** and 79 of which were labeled **Banded**. Experts can use this tree to predict the likelihood of banding on subsequent tuples.

For example, if the **chrome solution ratio** is **high**, **ink temperature** is **low**, and **viscosity** of the ink is **high**, then banding appears unlikely. In conjunction, these three conditions constitute a favorable condition for printing. In contrast, if a cylinder was plated using a **chrome solution ratio** that is **very low**, it appears that banding is very likely.

Since many of the measured attributes are continuously-valued, our TDIDT variation partitions the numeric attributes so that the resulting finite, discrete values best discriminate the possible outcomes. [\[link to Section C3.4\]](#). The strategy that we developed is a greedy, local (i.e., intervals developed for each decision-tree node), n-interval ($2 \leq n \leq 5$) divisive strategy. For purposes of illustration, we have used symbolic labels in place of discovered ranges in the tree of Figure 1.

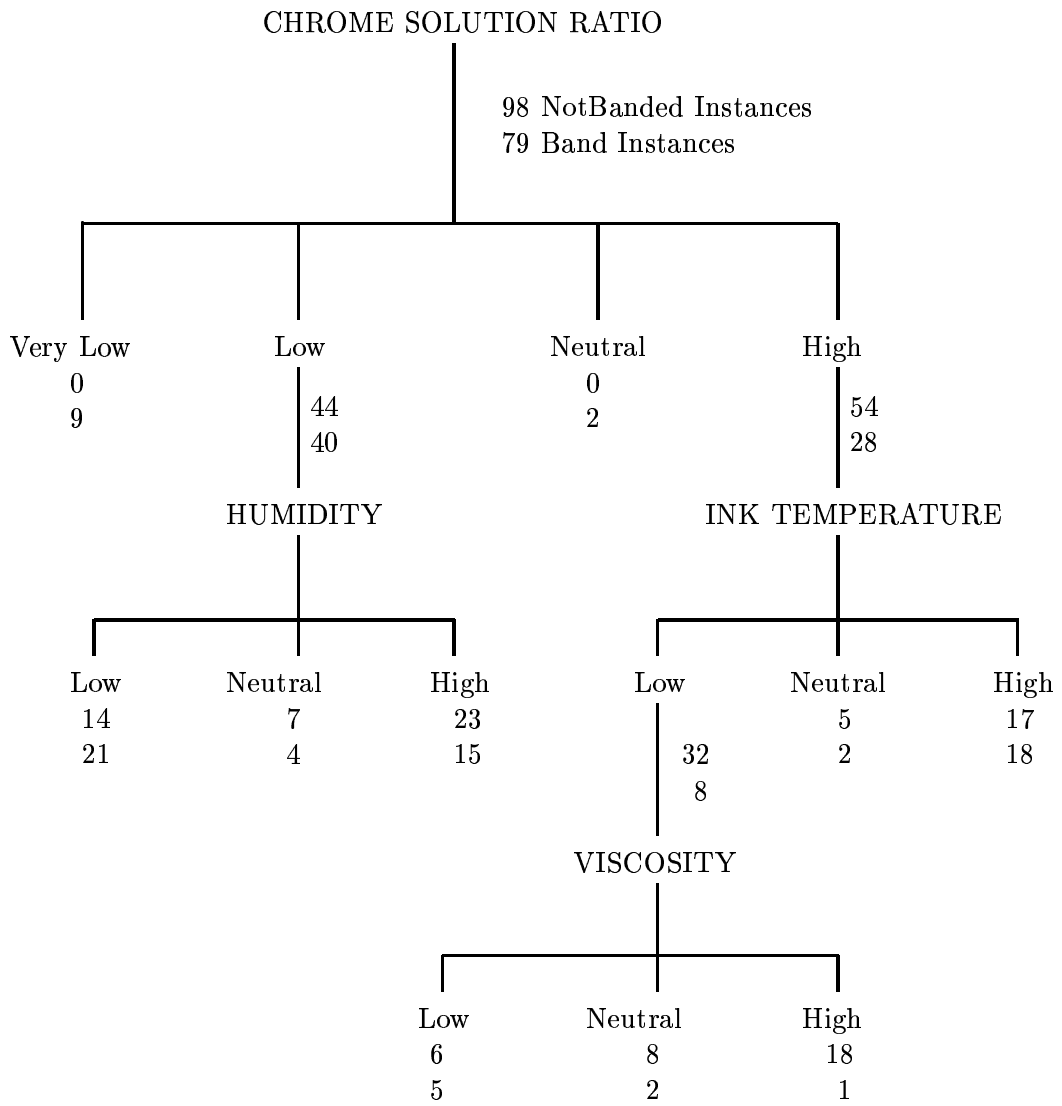


Figure 1: A sample decision tree.

H1.3.1.2 Interactive data mining using APOS

A TDIDT algorithm called APOSTM¹ (Evans and Fisher 1994) was one of three participants in the *interactive induction* process (Buntine and Sterling 1990). An analyst, Bob Evans, acted as an interface between TDIDT and an expert on press operations, though Evans, who has more than 35 years experience in the printing industry, sometimes played the dual role of analyst and expert.

In our setting, experts did not possess strong knowledge of banding causes, but were able to supply relevant attributes and offer hypotheses about why APOSTM indicated that certain attribute interactions (i.e., paths in the decision tree) resulted in high incidents of banding or not. The interaction between machine, expert, and analyst is an iterative process where the expert and analyst reflect on machine-induced rules. The expert's background knowledge may expedite induction under sparse data conditions by discounting attributes that would be discounted anyway given more data. This philosophy of using weak knowledge to filter inductive processing also lies behind work by Clark and Matwin (1993). If this strategy is followed naively, however, it may exclude the discovery of novel findings that violate expert expectations.

In general, interactive induction is part of a repertoire of tools for eliciting knowledge from experts, which follows in the tradition of other machine learning approaches to expert system construction (Michalski and Chilausky 1980; Bareiss, Porter, and Weir 1989). Induction over data enumerates rules that appear consistent with class labelings. The expert can often explain inductively-acquired rules, even if the rules were not known to the expert to begin with. In a system/analyst/expert partnership, the system greatly augments the processing and memory capacity of the expert, promoting an evolving conception of the problem in the expert's mind.

H1.3.1.3 Experiences and Results in Banding Analysis

We have described the methods that we chose to address many of our technological concerns in banding analysis, but our experience also suggests that social and cost issues all play a role in the effective acquisition and implementation of knowledge in an industrial setting. Based on our experience,

¹APOS is an acronym for A Posterior, and is a trademark of R R Donnelley and Sons.

we posit the general iterative model of Figure 2 as a guide to knowledge acquisition in an industrial setting.

The model suggests that there are many opportunities for feedback and revision of past decisions, but we have distinguished ‘rule implementation’ and ‘performance evaluation’ as falling outside the main interactive loop. By this, we stress that analysts and experts must take great care before implementing recommendations that stem from data mining into actual operations.

Problem Specification

Donnelley management initially had the modest goal of mitigating banding, as opposed to eliminating it completely, though of course the latter would be a desirable outcome. This implied that neither a *deterministic* or *causal* model of banding need be constructed. Rather, a statistically-reliable model would suffice. An inductive approach was chosen for this reason.

Method Selection

There are many forms of inductive reasoning. We initially chose decision tree induction because managers and craftsmen were familiar with decision trees. Moreover, the process of attribute selection in decision-tree construction was accessible to industry users. In contrast, neural network and statistical models [[link to Sections E1,E7](#)] (Weiss and Kulikowski 1991) may be understood by those trained in their proper use, but they may be less accessible to a wider population of professionals.

Data Definition and Engineering

The analyst and experts defined the attributes and class labels, which was not a trivial process. For example, some analysis of initial data produced the class labeling (and data exclusion) rules described earlier. Furthermore, the initial set of attributes was not the final set. Printing engineers initially identified several attributes as relevant only to discard them later based on TDIDT output. For example, the initial set included a subjective attribute called `chrome condition` with values `clear` and `cloudy`. No rules were found using this attribute because the attribute’s value could not be reliably determined, and experts eventually dropped the attribute from the data definition.

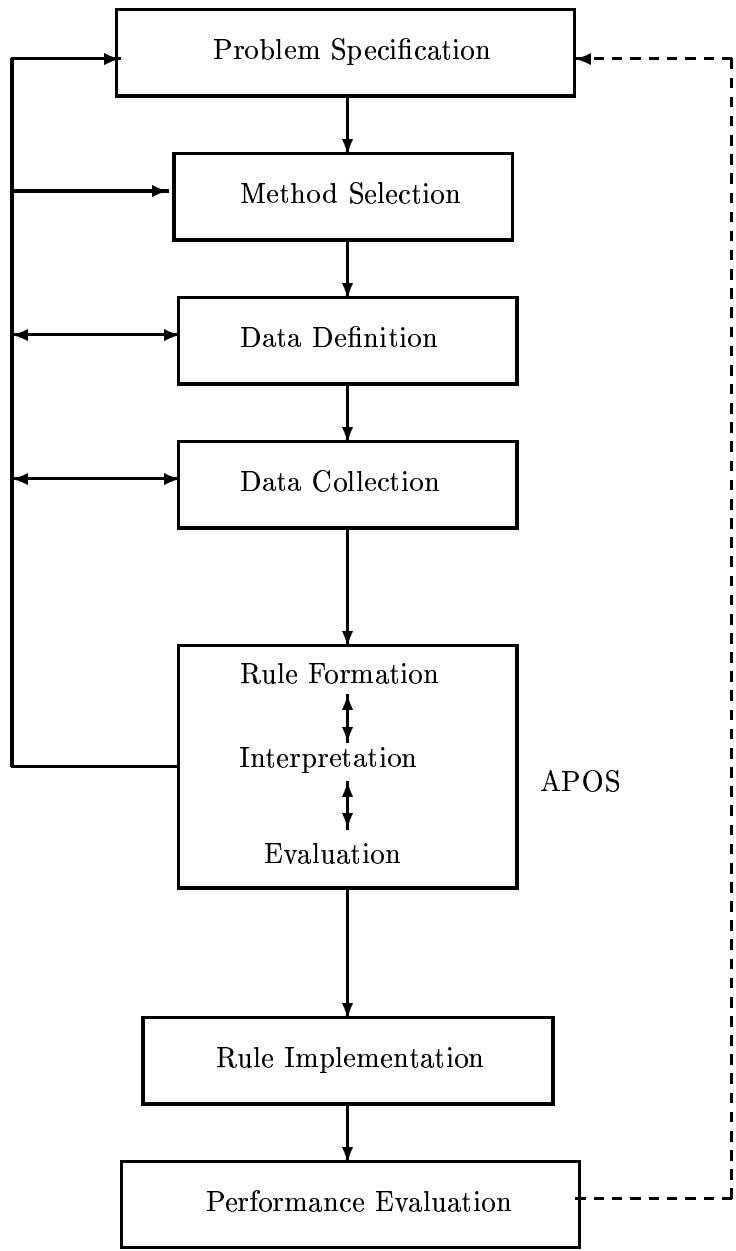


Figure 2: An industrial process improvement model.

Finally, in many domains inductive methods are good at identifying irrelevant attributes, but no inductive method can identify new primitive attributes. Under ideal conditions, it is safer to err by initially generating a liberal superset of the suspected relevant attributes.

Data Collection

Superficially, data collection may seem trivial, but together with data engineering it constituted the primary bottleneck for interactive induction in our setting. In the banding application, data collection was largely a manual process of completing data collection forms. Unfortunately, craftsmen faced time constraints that did not always allow accurate and complete data collection. The analyst acknowledged these difficulties, and stressed that when faced with conflicts, it is better for data to be incomplete (i.e., some attribute values not recorded) than to record data that was incorrect or even arbitrary.

Data collection may also introduce time conflicts that impact data definition. Although we noted under data definition and engineering that a large set of measurable attributes is desirable under ideal circumstances, measuring too many attributes can require too much time, thus motivating a judicious reduction in the number of attributes being measured.

Effective induction discriminates conditions that lead to differing outcomes; this requires that data be collected on examples of all the possible outcomes – `Banded` and `notBanded`. However, the common wisdom in our setting, and we suspect others, is that examples need only be collected when problems such as banding occur. In part, this tendency undoubtedly stems from the way that the problem is defined – to mitigate or eliminate banding – in which the stress is placed on the undesirable outcome. We revised the problem statement to that of increasing the occurrence of desirable outcomes (e.g., `notBanded`), which helped motivate craftsmen to collect `notBanded` data, as well as `Banded` data.

Rule Formation, Interpretation, and Evaluation

In an interactive induction setting the tasks of forming, interpreting, and evaluating rules are tightly coupled. Experts compensate for system brittleness in the face of sparse or biased data. In many cases, the expert may reject the system's principle recommendation. For example, in one session

APOSTM recommended that the data be broken up by press (eg., TR802, TR813) since bands seemed to occur more frequently on some presses over others. This line of reasoning may have led to some insights, but the expert preferred an alternative attribute that he felt would produce results more generally applicable to the printing process.

The analyst and the expert must further evaluate the final rule set. The tree of Figure 1 reveals one rule that isolates banding incidents (i.e., **chrome solution ratio very low**) and one rule strongly suggests no banding (i.e., **chrome solution ratio high, ink temperature low, and viscosity high**). Other rules (i.e., paths) stemming from the tree are ambiguous as to the outcome. In fact, the large number of ambiguous paths suggests that a system such as BRUTE (Riddle, Segal, and Etzioni 1994), which learns decision rules directly, rather than as paths in a decision tree, may be a better basis for interactive induction in a setting like ours. The decision tree learner produced many rules of no interest, because the learner was constrained to forming a tree.

In general, single applications of decision-tree induction uncovered few rules to reliably control cylinder banding. Fortunately, few rules are needed in a setting like ours, since craftsmen can control many attributes, thus effectively biasing the environment in response to even a few discovered rules. In practice, several experts applied APOSTM many times and collected approximately 15 rules in total from all of the sessions.

Rule Implementation

Discovered rules must be ‘implemented’. In the case of cylinder banding, press operators should control attributes (e.g., **ink temperature**) toward values leading to favorable outcomes (i.e., **notBanded**) and/or away from values leading to undesirable outcomes (i.e., **Banded**). Induction results in rules of theoretical merit (i.e., following their recommendations will mitigate banding), but experts had to evaluate rules by pragmatic considerations such as *cost-effectiveness* and *social acceptance*. Several rules were discarded on pragmatic grounds, reducing the rule set from approximately 15 to 12.

For example, some rules suggested that humidifying the plant would mitigate banding, but management discarded this recommendation on an expense basis. In effect, management designated **humidity** an uncontrollable parameter. Nonetheless, it still may be useful to have such attributes included in

rules, because they define differing contexts under which differing actions are in order. In other situations, a rule may indicate an attribute that can be controlled, but to do so would have undesirable side effects. The experts discarded at least one option found to mitigate cylinder banding on the basis of its potential negative impact on print quality along dimensions unrelated to banding.

There are also social issues to be considered. In particular, a table of rules of the form

```
chrome solution ratio is high, and
ink temperature is low, and
viscosity is high
  IMPLIES the cylinder will be notBanded.
```

presents several difficulties. First, individual craftsmen generally control only some of the attributes that participate in a rule. A Plater controls the **chrome solution**, whereas a Pressman controls **ink temperature** and **viscosity**. Rules that conjoin attributes across several areas of responsibility may, at least in the eyes of craftsmen and managers, make *blame assignment* difficult and ambiguous. Even within a press team, different team members might be responsible for controlling different parameters. Given the interactions between attributes within and across conjunctive rules, this can make management of press-wide parameters untenable.

For these reasons, an expert and analyst translated the set of conjunctive rules into a list of independent attribute values (or ranges) like those shown in Table 1 using a straightforward process that is not elaborated here. In retrospect, this final choice of representation might motivate a different discovery method than decision tree induction, but a naive approach that discovers simple relationships between single attributes and outcomes directly, would not implicate as many informative attributes as an approach that takes an intermediate step of discovering interacting attributes that collectively inform outcome (e.g., decision tree induction).

The experts posted a complete list of conditions like those shown in Table 1 for the press operators and the plating craftsmen. The printing experts told the craftsmen to stay within the favorable value ranges unless quality considerations dictated otherwise. The experts also explained that, based on the APOSTM generated trees and the expert's own past experience, operating

Table 1: Recommendations translated from APOS rules.

attribute	FAVORABLE	NEUTRAL	UNFAVORABLE
Humidity	High	Medium	Low
Grain Screen	Yes	–	No
Anode Distance	Wide	Medium	Narrow
Chrome Ratio	High	Medium	Low
Ink Temperature	Low	Medium	High
Ink Viscosity	High	Medium	Low
Blade Stroke	Long	Medium	Short
Current Density	High	Medium	Low

Table 2: Absolute number of bands.

1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
538	384	138	66	42	109	21	26	37	26

with two or more attribute values within the unfavorable ranges would greatly increase the risk of banding.

Performance Evaluation

Table 2 graphs the absolute number of bands per year since 1989. The trends are almost identical when the data is expressed in terms of the relative number of bands (per unit work).

The peak during 1994 occurred in September of that year, after contaminated, abrasive ink entered the system in late August. Ink was implicated by rules generated by APOSTM early in the study. Technicians purged the ink storage facility and cylinder band performance returned to 1993 levels.

Unfortunately, an industrial environment does not facilitate the kind of experimentation that occurs in a laboratory environment. External factors, which may be excluded or at least controlled in a laboratory environment, can impact an industrial operation and preempt precise assignment of cause and effect. One confound in our application was the input of an industrial consultant, who provided some recommendations in 1989. We compared the recommendations of the consultant with those of APOSTM, and found that

APOSTM produced significantly more rules (12-15) than the consultant (3), which filled important gaps in the consultant's recommendations.

Moreover, APOSTM also recommended more conservative thresholds on numeric variables in the case where the the system and consultant recommendation qualitatively matched. In general, APOSTM recommendations were specific to the plant from which data was collected, whereas the consultant's recommendations were more general. Finally, one APOSTM-generated recommendation conflicted with one consultant recommendation. Local engineers performed a deeper analysis and agreed with the APOSTM rule, though the APOSTM rule initially seemed counterintuitive.

Our analysis convinces us that APOSTM-generated rules are primarily responsible for post-1990 improvements: APOSTM-generated rules subsume, and in one case contradict, those of the consultant, and in any case they are the sole basis for current operations.

Despite our success, rule implementation should proceed cautiously. For example, the experts were very careful to approach a printing crew that seemed most open to input. Even after the experts distributed the APOSTM recommendations, some time elapsed before all crews were following the recommendations. The continued improvement illustrated by Table 2 is not due to the generation of new rules, but primarily to the craftsmen gradually accepting the recommendations informally distributed in December, 1990 and formally distributed in April, 1991.

Once recommendations are made and they begin affecting operations, it is very difficult to turn back. If recommendations have a negative effect, then confidence in the knowledge acquisition process will be seriously compromised. In our case, rules were not implemented unless they were justified and explained by plant engineers, but a more cautious strategy might have deferred implementation until these rules were systematically tested on subsequent data. In contrast, if implemented rules have a positive effect, then the sample space is inalterably changed, since management is unlikely to return to the old situation for purposes of hypothesis testing.

H1.3.1.4 Concluding Remarks

Though our application did not involve massive data sets (i.e., we worked with about 500 tuples), the interactive data mining approach still significantly extended the processing capabilities of human analysts. Moreover, an

interactive data mining approach *demand*s and (management willing) institutionalizes a disciplined approach to problem solving, which includes the systematic collection and analysis of data. Even if a data mining technique were not theoretically required to discover important attribute interactions in the data from a data processing standpoint, it is likely that the long-standing banding problem would not have been mitigated to the extent that it was without the introduction of an induction technique.

We have noted that social factors played a significant role in the success of our project. Along these lines it seemed important that a reward system be in place to promote acceptance of induced rules. At the Donnelley plant this reward system relied on the craftsmen's desire to efficiently produce high-quality products and to receive recognition when they reach remarkable quality or production milestones. In particular, pictures and congratulatory comments were posted on a bulletin board outside the cafeteria about crews who have achieved a year or more of production without any delays due to cylinder bands. Though records are incomplete, it appears that no crew in the plant had ever achieved one year without a cylinder band prior to the application of data mining to the problem. More importantly, the crews regard the accomplishment as noteworthy and the crew members themselves have taken on the role of educating new members on the merits of system-translated recommendations.

[Word Count: 3198]

References

- Bareiss, R., B. Porter, and K. Murray. 1989. "Supporting start-to-finish development of knowledge bases." *Machine Learning* 4: 259–284.
- Buntine, W. and D. Sterling. 1990. "Interactive induction." In *Machine Intelligence* 12, edited by J. E. Hayes-Michie, D. Michie, and E. Tyugu, pp. 121-138. Oxford England: Oxford University Press.
- Clark, P. and S. Matwin. 1993. "Using qualitative models to guide inductive learning." In *Proceedings of the 1993 International Machine Learning Conference (Amherst, MA)*, edited by P. Utgoff, pp. 49-56. San Francisco, CA: Morgan Kaufmann.

Evans, B. and D. Fisher. 1994. "Overcoming process delays with decision tree induction." *IEEE Expert* **9**: 60–66.

Michalski, R. and C. Chilausky. 1980. "Learning by being told and learning from examples: An experimental comparison of the two methods of knowledge acquisition in the context of developing an expert system for soybean disease diagnosis." *International Journal Policy Analysis and Information Systems* **4**: 125–161.

Quinlan, J. R. 1991. *C4.5: Programs for Machine Learning*. San Francisco, CA: Morgan Kaufmann.

Riddle, P., R. Segal, and O. Etzioni. 1994. "Representation design and brute-force induction in a Boeing manufacturing domain." *Applied Artificial Intelligence* **8**: 125–147.

Weiss, S. and C. Kulikowski. 1991. *Computer Systems That Learn*. San Francisco, CA: Morgan Kaufmann.