

View Reviews

Paper ID

4143

Paper Title

Algebraic Ground Truth Inference for Possibly Dependent Binary Classifiers via Data Moments

Reviewer #2

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words (1-2 sentences or paragraphs).

This paper provides an algebraic study of the problem of ground truth inference from an ensemble of binary classifiers. It is shown that the ground truth label statistics together with the classifier accuracies are encoded as roots in a polynomial system of equations, parametrized by the decision sample statistics of the classifiers. Of special interest is the case of three binary classifiers. If these classifiers are independent the polynomial system consists of 7 quartic non-homogeneous equations in 7 unknowns. A Groebner basis computation shows that this is a zero-dimensional system, that is, it admits a finite number of solutions, and in particular two solutions, one of which are the ground truth unknown statistics. The nature of the solution provides algebraic tests for checking the independence of classifiers and this scheme is explored to the case of 4 possibly weakly dependent classifiers. The paper also contains an experimental evaluation for this latter case.

2. [Detailed comments] Describe the strengths and weaknesses of the work, with respect to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, relation with prior work, clarity of writing, and relevance to the ICML community.

STRENGTH 1: This is a quite interesting paper of theoretical flavor, with potential practical implications.

STRENGTH 2: The paper is mostly well written in a somewhat casual style that is pleasant to follow and rather adds than subtracts from its value.

STRENGTH 3: The use of algebraic geometry is delivered in a manner friendly to the non-expert reader.

WEAKNESS 1: A little more rigor is needed. In particular,

- i) please formally define the prevalences and accuracies ϕ . Currently this is only done in the appendix.
- ii) The statement of Theorem 3 is not precise. What is shown is that for independent classifiers the root to the quadric is rational. But as noted in the text itself, dependent classifiers may also give a rational root.
- iii) Please be precise and consistent when using the term "independent". Right now we have "independent polynomials" (205, column 1) "independent solution quadratic polynomial" (262, column 1), "independent polynomial system" (240, column 2), "independent errors algebraic solution" (284, column 2). I suggest coining an elegant terminology for indicating whether an object (system, polynomial, solution, e.t.c.) is associated to independent classifiers and using that throughout. Note that in algebraic geometry we have the different notion of algebraically independent polynomials, with which confusion should be avoided.
- iv) in 240, column 2 please express more rigorously "can be fed back into the ϕ_α quadric". In fact, this expression only makes sense because the f quantities are expressible as functions of the ϕ quantities, perhaps remind the reader of this.
- v) in 296, column 1, please be more precise and say that the Buchberger algorithm is in general exponential in the number of variables and the maximum degree of the polynomial equations, rather than saying "exponential in space and time".

vi) please mention the term order that is used for the Groebner basis computation and comment on whether this has an effect.

WEAKNESS 2: The description of the error minimization approach in 294-303, column 2 is unclear.

WEAKNESS 3: There are some minor language issues, please have the paper proofread by a proficient user of English. For example in 86, column 2 the phrase "We do not mean to imply to continuing to point out there is no probability in the algebraic approach that the above approaches are inferior." should be revisited. Also, syntax issues in 19, column 2 and 28-30, column 2.

SUGGESTION: Add a little proof for equations (1) and (2).

I have two concerns:

CONCERN 1: As mentioned in Platanios 2016, in practice, binary classifiers are rarely independent. This fact seems to be reducing the significance of the paper. How do the authors respond to that?

CONCERN 2: In 315-317 the authors say that they resolve the ambiguity between the two solutions by assuming knowledge of whether the prevalence of alpha is greater or less than 50%. Does that not render the experimental comparison with other methods somewhat unfair?

%%%%%%%%%

POST REBUTTAL COMMENTS

Since the authors did not respond to my two concerns, I am forced to reduce my score.

3. [Relevance and Significance] (Is the subject matter important? Does the problem it tries to address have broad interests to the ICML audience or has impact in a certain special area? Is the proposed technique important, and will this work influence future development?)

Solid contribution to relevant problem

4. [Novelty] (Is relation to prior work well-explained, does it present a new concept or idea, does it improve the existing methods, or extend the applications of existing practice?)

One idea that surprised me by its originality, solid contributions otherwise

5. [Technical quality] (Is the approach technically sound. The claims and conclusions are supported by flawless arguments. Proofs are correct, formulas are correct, there are no hidden assumptions.)

Technically adequate for its area, solid results

6. [Experimental evaluation] (Are the experiments well designed, sufficient, clearly described? The experiments should demonstrate that the method works under the assumed conditions, probe a variety of aspects of the novel methods or ideas, not just the output performance, present comparisons with prior work, test the limits and check the robustness of the novel methods or ideas, and demonstrate their practical relevance.)

Sufficient evaluation w.r.t. most criteria

7. [Clarity] (Is the paper well-organized and clearly written, should there be additional explanations or illustrations?)

Mostly clear, but improvements needed, as recommended in the detailed comments.

8. [Reproducibility] (are there enough details to reproduce the major results of this work?)

No

10. Please provide an "overall score" for this submission.

Weak Reject: Borderline, tending to reject

11. [Confidence] Please provide your confidence in your assessment of this submission.

I am an expert of the area and very confident about my evaluation.

16. Please acknowledge that you have read the author rebuttal. If your opinion has changed, please summarize the main reasons in the Detailed comments sections.

Agreement accepted

Reviewer #4

Questions**1. [Summary] Please summarize the main claims/contributions of the paper in your own words (1-2 sentences or paragraphs).**

In this paper, a new point of view is put forward to solve the problem of truth inference, that is, using algebraic statistics method to infer the truth. The proposed method uses three classifiers to complete the inference, and uses the fourth classifier to judge whether the voting classifiers are independent of each other. Experimental results show that the proposed method outperforms the traditional MV algorithm on several datasets.

2. [Detailed comments] Describe the strengths and weaknesses of the work, with respect to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, relation with prior work, clarity of writing, and relevance to the ICML community.

In this paper, a new point of view is put forward to solve the problem of truth inference, that is, using algebraic statistics method to infer the truth. The proposed method uses three classifiers to complete the inference, and uses the fourth classifier to judge whether the voting classifiers are independent of each other. Experimental results show that the proposed method outperforms the traditional MV algorithm on several datasets.

In this paper, it is a novel idea to deal with inference by using algebraic statistic, which is method is different from the common probability graph model. In the introduction, the author listed the advantages of this method which is not based on any hypothesis and data distribution. Another strength of the paper is that three theorems are provided to show the characteristics of the proposed method. The supplementary document provides the derivation in detail.

There are also some issues in the paper. In the introduction, when related work is reviewed, some of them are proposed in the context of crowdsourcing. Those truth inference methods do not require classifiers (or they treat humans as classifiers.) However, in the experiment, the authors used NN, SVM, Decision Tree, and Naïve Bayes as classifiers. Is that to say the proposed method relies on classifiers?

Another limitation of the proposed method is that it can only deal with binary classification. However, many practical problems are multi-class ones, and it is not clear whether this method is easy to expand to such scenarios.

The experimental design is not sufficient. In addition to the small number of validation data sets, the comparison method is only traditional MV. Truth inference has been studied for many years, without the comparison with more state-of-the-art methods, it is hard to say this new method is better in practice.

3. [Relevance and Significance] (Is the subject matter important? Does the problem it tries to address have broad interests to the ICML audience or has impact in a certain special area? Is the proposed technique important, and will this work influence future development?)

Solid contribution to relevant problem

4. [Novelty] (Is relation to prior work well-explained, does it present a new concept or idea, does it improve the existing methods, or extend the applications of existing practice?)

One idea that surprised me by its originality, solid contributions otherwise

5. [Technical quality] (Is the approach technically sound. The claims and conclusions are supported by flawless arguments. Proofs are correct, formulas are correct, there are no hidden assumptions.)

Technically adequate for its area, solid results

6. [Experimental evaluation] (Are the experiments well designed, sufficient, clearly described? The experiments should demonstrate that the method works under the assumed conditions, probe a variety of aspects of the novel methods or ideas, not just the output performance, present comparisons with prior work, test the limits and check the robustness of the novel methods or ideas, and demonstrate their practical relevance.)

Insufficient, unfair, inadequate or uninformative comparisons with related work

7. [Clarity] (Is the paper well-organized and clearly written, should there be additional explanations or illustrations?)

Mostly clear, but improvements needed, as recommended in the detailed comments.

8. [Reproducibility] (are there enough details to reproduce the major results of this work?)

Yes

10. Please provide an "overall score" for this submission.

Weak Reject: Borderline, tending to reject

11. [Confidence] Please provide your confidence in your assessment of this submission.

I am quite sure about my evaluation. It's unlikely, although possible that I missed something that should affect my ratings.

Reviewer #5

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words (1-2 sentences or paragraphs).

In this paper, an exact polynomial system is proposed, the utility of the proposed algebraic approach is demonstrated.

2. [Detailed comments] Describe the strengths and weaknesses of the work, with respect to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, relation with prior work, clarity of writing, and relevance to the ICML community.

An exact polynomial system is proposed in this article. However, the innovation of this article is limited. The method of how to reduce the cost of misclassification hasn't been elaborated clearly. The effect of the proposed method should be further proved. By the way, the main contribution must be clearly stated. The author should pay more attention to the details as the paper should be formatted better and some grammar should be checked carefully, making sure that all parameters and symbols appear in this article have been explained.

The statement of Theorem 3 is not precise.

Please be precise and consistent when using the term "independent"

The proposed method can only deal with binary classification.

3. [Relevance and Significance] (Is the subject matter important? Does the problem it tries to address have broad interests to the ICML audience or has impact in a certain special area? Is the proposed technique important, and will this work influence future development?)

Reasonable contribution to a minor problem

4. [Novelty] (Is relation to prior work well-explained, does it present a new concept or idea, does it improve the existing methods, or extend the applications of existing practice?)

One idea that surprised me by its originality, solid contributions otherwise

5. [Technical quality] (Is the approach technically sound. The claims and conclusions are supported by flawless arguments. Proofs are correct, formulas are correct, there are no hidden assumptions.)

Technically adequate for its area, solid results

6. [Experimental evaluation] (Are the experiments well designed, sufficient, clearly described? The experiments should demonstrate that the method works under the assumed conditions, probe a variety of aspects of the novel methods or ideas, not just the output performance, present comparisons with prior work, test the limits and check the robustness of the novel methods or ideas, and demonstrate their practical relevance.)

Sufficient evaluation w.r.t. most criteria

7. [Clarity] (Is the paper well-organized and clearly written, should there be additional explanations or illustrations?)

Mostly clear, but improvements needed, as recommended in the detailed comments.

8. [Reproducibility] (are there enough details to reproduce the major results of this work?)

Yes

9. [Questions for authors] Please provide questions for authors to address during the author feedback period. (Optional, to help authors focus their response to your review.)

1. The effect of the proposed method should be further proved. A comparison between the proposed method and recent research findings will be helpful.

2. The main contribution must be clearly stated.

3. The paper should be formatted better and some grammar should be checked carefully.

4. Author should make sure that all parameters and symbols appear in this article have been explained.

10. Please provide an "overall score" for this submission.

Weak Reject: Borderline, tending to reject

11. [Confidence] Please provide your confidence in your assessment of this submission.

I am knowledgeable and willing to defend my evaluation, but there's a chance I missed something.

16. Please acknowledge that you have read the author rebuttal. If your opinion has changed, please summarize the main reasons in the Detailed comments sections.

Agreement accepted

Reviewer #6

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words (1-2 sentences or paragraphs).

For the ground truth information problem, this paper constructs an algebraic method based on sample statistics. It proves that if sample errors of the classifiers are independent, only three classifiers are needed to solve the algebraic system.

2. [Detailed comments] Describe the strengths and weaknesses of the work, with respect to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, relation with prior work, clarity of writing, and relevance to the ICML community.

The article is relatively difficult to understand, many concepts in use are not explained in detail. The frequency mentioned in Theorem 1 is a measured value, which in fact is not completely accurate, which would affect the accuracy of the system. In addition, the claim in section 4 "this system is identifiable" is actually not convincing. The author should at least give the explanation. It is difficult to generate independent classifiers in the experimental part.

3. [Relevance and Significance] (Is the subject matter important? Does the problem it tries to address have

broad interests to the ICML audience or has impact in a certain special area? Is the proposed technique important, and will this work influence future development?)

Reasonable contribution to a minor problem

4. [Novelty] (Is relation to prior work well-explained, does it present a new concept or idea, does it improve the existing methods, or extend the applications of existing practice?)

Minor variations to existing techniques

5. [Technical quality] (Is the approach technically sound. The claims and conclusions are supported by flawless arguments. Proofs are correct, formulas are correct, there are no hidden assumptions.)

Claims not completely supported, assumptions or simplifications unrealistic

6. [Experimental evaluation] (Are the experiments well designed, sufficient, clearly described? The experiments should demonstrate that the method works under the assumed conditions, probe a variety of aspects of the novel methods or ideas, not just the output performance, present comparisons with prior work, test the limits and check the robustness of the novel methods or ideas, and demonstrate their practical relevance.)

Insufficient, unfair, inadequate or uninformative comparisons with related work

7. [Clarity] (Is the paper well-organized and clearly written, should there be additional explanations or illustrations?)

This paper is difficult to understand in places because of typos, lack of organization, or another flaws. The paper would need significant/major improvement in writing.

8. [Reproducibility] (are there enough details to reproduce the major results of this work?)

No

10. Please provide an "overall score" for this submission.

Weak Reject: Borderline, tending to reject

11. [Confidence] Please provide your confidence in your assessment of this submission.

I am knowledgeable about this area, but I am not certain about, and am willing to change my evaluation.