The Research Gamble

The standard accuracy measures used to evaluate neural nets, learning, parsing, tagging, searching, etc. come from the search application where we have a pool of documents, the corpus, some of which are relevant (+R) but most of which are not (-R). The search procedure returns a set of documents predicted as useful (+P) and others that are predicted as irrelevant (-P).

The proportion of relevant documents returned is recall, $\frac{|+P|}{|R|}$, whereas the proportion of returned documents relevant is precision, $\frac{|+P|}{|+P+R|}$. The model predicted 40 out of 60 cases and 60 – ve cases, but in this example we were just guessing so we expect 40% recall just because we have 40% positive labels.

The problem is that keyword search returns masses of “hits” and we need to be valid. Multiplying the figures by some large factor whilst retaining the probability or proportion of time that we are making an informed decision shows that fair bookmaker odds define the unique formula for determining the percentage return. We will work with fair odds based on statistics on past performance. In our example the horse has won 30 out of 100 starts leading to odds of 7:3 for our field. This means if the horse wins I win 70% or a $7 bet that I stand to lose if it loses.

In defining Bookmaker formally, we make up of sample probabilities from the continuous matrix, $P_{x,y} = \Phi(x + y)$ and $\Phi(x) = \frac{1}{1 + e^{-x}}$ and also $P_{x,y} = \Phi(x) \Phi(y) + \Phi(y) \Phi(x)$.

Trotting out Recall and Precision

In gambling, the house always wins, and in horseracing, the bookmaker is no exception. The basis of the odds is set by a bookie is the assessed likelihood of a horse winning, as influenced by talk and bets from those in the know. The bookie will then add on a percentage as she calculates the odds. This bookie doesn’t know anything – he’s just guessing! A real bookie would reflect sights of his money in no time. But the fair bookmaker algorithm simply reports that his edge is zero.

The Bookmaker always Wins

For information retrieval, machine learning and neural nets, the problem is not normally too little data but too much. Large amounts of data are used in training and significant proportions are set aside for validation and testing to avoid overtraining. In information retrieval or web search, the problem is that keyword search returns masses of “hits” and we need to assess how useful the results are – that is we want to know the “accuracy”.

Recall and Precision are Losers

Recall and precision suffer from a number of disadvantages that make them unsuitable for defining an accuracy measure:

- They assess only a single condition
- There is a tradeoff between them
- Neither can be interpreted alone
- They ignore the cost of errors
- Each is easily inflated.

Recall by labeling more cases we Precision by labeling hard ones –ve Accuracy can be Biased

The weighted averages of correct cases of recall and inverse precision, and inverse precision, are all equivalent accuracy measures. But this definition does not take into account the cost of errors or the baseline for guessing.

The Bookmaker always Wins

In gambling, the house always wins, and in horseracing, the bookmaker is no exception. The basis of the odds is set by a bookie is the assessed likelihood of a horse winning, as influenced by talk and bets from those in the know. The bookie will then add on a percentage as she calculates the odds. This bookie doesn’t know anything – he’s just guessing! A real bookie would reflect sights of his money in no time. But the fair bookmaker algorithm simply reports that his edge is zero.

The Bookmaker always Wins

For information retrieval, machine learning and neural nets, the problem is not normally too little data but too much. Large amounts of data are used in training and significant proportions are set aside for validation and testing to avoid overtraining. In information retrieval or web search, the problem is that keyword search returns masses of “hits” and we need to assess how useful the results are – that is we want to know the “accuracy”.

Recall and Precision are Losers

Recall and precision suffer from a number of disadvantages that make them unsuitable for defining an accuracy measure:

- They assess only a single condition
- There is a tradeoff between them
- Neither can be interpreted alone
- They ignore the cost of errors
- Each is easily inflated.

Recall by labeling more cases we Precision by labeling hard ones –ve Accuracy can be Biased

The weighted averages of correct cases of recall and inverse precision, and inverse precision, are all equivalent accuracy measures. But this definition does not take into account the cost of errors or the baseline for guessing.

The Bookmaker always Wins

In gambling, the house always wins, and in horseracing, the bookmaker is no exception. The basis of the odds is set by a bookie is the assessed likelihood of a horse winning, as influenced by talk and bets from those in the know. The bookie will then add on a percentage as she calculates the odds. This bookie doesn’t know anything – he’s just guessing! A real bookie would reflect sights of his money in no time. But the fair bookmaker algorithm simply reports that his edge is zero.

The Bookmaker always Wins

For information retrieval, machine learning and neural nets, the problem is not normally too little data but too much. Large amounts of data are used in training and significant proportions are set aside for validation and testing to avoid overtraining. In information retrieval or web search, the problem is that keyword search returns masses of “hits” and we need to assess how useful the results are – that is we want to know the “accuracy”.

Recall and Precision are Losers

Recall and precision suffer from a number of disadvantages that make them unsuitable for defining an accuracy measure:

- They assess only a single condition
- There is a tradeoff between them
- Neither can be interpreted alone
- They ignore the cost of errors
- Each is easily inflated.

Recall by labeling more cases we Precision by labeling hard ones –ve Accuracy can be Biased

The weighted averages of correct cases of recall and inverse precision, and inverse precision, are all equivalent accuracy measures. But this definition does not take into account the cost of errors or the baseline for guessing.

The Bookmaker always Wins

In gambling, the house always wins, and in horseracing, the bookmaker is no exception. The basis of the odds is set by a bookie is the assessed likelihood of a horse winning, as influenced by talk and bets from those in the know. The bookie will then add on a percentage as she calculates the odds. This bookie doesn’t know anything – he’s just guessing! A real bookie would reflect sights of his money in no time. But the fair bookmaker algorithm simply reports that his edge is zero.

The Bookmaker always Wins

For information retrieval, machine learning and neural nets, the problem is not normally too little data but too much. Large amounts of data are used in training and significant proportions are set aside for validation and testing to avoid overtraining. In information retrieval or web search, the problem is that keyword search returns masses of “hits” and we need to assess how useful the results are – that is we want to know the “accuracy”.

Recall and Precision are Losers

Recall and precision suffer from a number of disadvantages that make them unsuitable for defining an accuracy measure:

- They assess only a single condition
- There is a tradeoff between them
- Neither can be interpreted alone
- They ignore the cost of errors
- Each is easily inflated.

Recall by labeling more cases we Precision by labeling hard ones –ve Accuracy can be Biased

The weighted averages of correct cases of recall and inverse precision, and inverse precision, are all equivalent accuracy measures. But this definition does not take into account the cost of errors or the baseline for guessing.

The Bookmaker always Wins

In gambling, the house always wins, and in horseracing, the bookmaker is no exception. The basis of the odds is set by a bookie is the assessed likelihood of a horse winning, as influenced by talk and bets from those in the know. The bookie will then add on a percentage as she calculates the odds. This bookie doesn’t know anything – he’s just guessing! A real bookie would reflect sights of his money in no time. But the fair bookmaker algorithm simply reports that his edge is zero.

The Bookmaker always Wins

For information retrieval, machine learning and neural nets, the problem is not normally too little data but too much. Large amounts of data are used in training and significant proportions are set aside for validation and testing to avoid overtraining. In information retrieval or web search, the problem is that keyword search returns masses of “hits” and we need to assess how useful the results are – that is we want to know the “accuracy”.

Recall and Precision are Losers

Recall and precision suffer from a number of disadvantages that make them unsuitable for defining an accuracy measure:

- They assess only a single condition
- There is a tradeoff between them
- Neither can be interpreted alone
- They ignore the cost of errors
- Each is easily inflated.

Recall by labeling more cases we Precision by labeling hard ones –ve Accuracy can be Biased

The weighted averages of correct cases of recall and inverse precision, and inverse precision, are all equivalent accuracy measures. But this definition does not take into account the cost of errors or the baseline for guessing.