**ON THE VALUE OF STOCHASTIC SIDE INFORMATION IN ONLINE LEARNING**

*Junzhang Jia, Xuetong Wu, Jamie Evans, Jingge Zhu*

University of Melbourne

Department of Electrical and Electronic Engineering Parkville, Victoria, Australia

# ABSTRACT

We study the effectiveness of stochastic side information in de- terministic online learning scenarios. We propose a forecaster to predict a deterministic sequence where its performance is evalu- ated against an expert class. We assume that certain stochastic side information is available to the forecaster but not the experts. We de- fine the minimax expected regret for evaluating the forecaster’s per- formance, for which we obtain both upper and lower bounds. Con- sequently, our results characterize the improvement in the regret due to the stochastic side information. Compared with the classical

online learning problem with regret scales with *O*(√*n*), the regret

can be negative when the stochastic side information is more pow- erful than the experts. To illustrate, we apply the proposed bounds to two concrete examples of different types of side information.

***Index Terms*—** Online learning, Expert advice, Minimax regret, Side information

# INTRODUCTION

The online learning problem aims to make predictions for prob- abilistic/deterministic instances which arrive sequentially, and has become significantly popular in game theory and learning theory fields recently. Merhav and Feder [1] studied online learn- ing problems for stochastic setup from an information-theoretic perspective, followed by [2]. In the deterministic setting, we will usually introduce a class of competitive predictors providing advice to the forecaster, namely the expert class [1, 3], and the learning performance is evaluated by the regret, i.e. the loss gap between the proposed forecaster and the best expert. To effec- tively leverage the experts, Littlestone and Warmuth [4] proposed a weighted majority algorithm, and the follow-up works such as [3, 5, 6] further proposed the randomized algorithms which produce logarithmic regret. In a more specific setup, Haussler et al. [7] considered binary and continuous instance spaces and provided an Ω( *n*log*N*) worst-case regret, where *n* is the sample size and *N* is the number of experts. With respect to different loss functions, Cesa-Bianchi and Lugosi [8] and Vanli and Kozat

√

[9] provided explicit upper and lower bounds on the regret for absolute loss and squared loss, respectively.

As a common situation in practice, the forecaster could access some additional resources which we call it *side information*, that

may provide some useful knowledge on the sequence of interest. Cover and Ordentlich [10] first studied a portfolio investment problem where the sequence of interest is the stock vectors that may depend on some finite-valued states (as side information), and their proposed forecaster can achieve the same wealth as the best side information dependent investment strategy. Xie and Barron

[11] studied the case when the sequence of interest is generated according to a pair-wise parametric distribution conditioning on the side information, and derived an logarithmic upper bound of the minimax regret. Cesa-Bianchi and Lugosi [8] analyzed the problem with an additional (deterministic) side sequence, then the learning performance depends on the occurrences of its agreed symbols compared to the sequence of interest. Recently, Bhatt and Kim [12] studied the probabilistic online learning problem where the side information is the auxiliary random symbols generated jointly with the data instance to be predicted, and analyzed the minimax regret under the logarithmic loss.

However, to the best of our knowledge there is no prior work discussing the formulation and effects of the stochastic side information under a deterministic online learning scenario. Inspired by the transfer learning problem [13] where people transfer the knowledge from one domain (source) to the domain of interest (target) with both the source and target data drawn from different but related distributions, we specify the formation of the side information that may depend on the target sequence with some stochasticity. In a similar spirit, we aim to explore the influence of a stochastic *sequential side information* (SSI) for predicting a *target sequence* of interest. In this paper, we propose a novel problem formulation where a forecaster tries to predicts a deterministic sequence with some stochastic side information, which is not known to the expert class. Then we develop an online learning framework with the expert class where we will additionally leverage the side information for prediction to min- imise the regret with respect to the best expert. With the proposed algorithm, we provide both the lower and upper bounds on the minimax regret under the absolute loss, where the target sequence is selected adversarially to maximise the regret. From the results, we show that introducing SSI can improve the typical learning rate in [1, 7, 8] if the side information performs better than the best expert. On the other hand, the side information will not hurt our prediction if it fails to provide much useful information.

# PROBLEM FORMULATION AND MAIN RESULTS

* 1. **Prediction with Experts and Stochastic Side Information**

specified, the expectation is always taken over the SSI conditional distribution *Pn*(S*n*|T*n*).

*R*(*n*):=infsup E *{L*(*T*˜*n*(*Sn,Tn*)*,Tn*)*−*min*L*(*Fθ,Tn*)*}.* (3)

*T*˜*n*

*n*

*Tn Sn θ*

We consider the online learning problem for a deterministic target sequence with the side information: we aim to design a

To evaluate the usefulness of the SSI, we introduce the maximum likelihood estimation of the target instances *XT* given *XS*:

forecaster that sequentially predict the outcome of an unknown *t t*

target sequence T*n* =(*XT ,XT ,...,XT* ) where each instance *XT X*ˆ *T* (*XS*)=argmax *P* (*XS*|*XT* )*.* (4)

1

2

*n*

*t*

*t*

*t*

*t*

*t*

takes value in a set X ⊆ R. The prediction of the forecaster at time *t*, denoted by *X*˜*t*, takes value in a space D which is a convex

*X*

and nonempty subset of R, and we also assume X ⊂D. We will

*T*

*t*

Then we denote the maximum likelihood prediction sequence by

Tˆ*n*(S*n*):=(*X*ˆ *T* (*XS*)*,X*ˆ *T* (*XS*)*,...,X*ˆ *T* (*XS*)). Furthermore, we

compare the forecaster with a class of experts. We use F*θ* to

*t* 1 *t* 2 *t n*

*n*

denote the prediction sequence made by the expert *θ*, and *fθ* to

make the assumption that the expected cumulative loss induced

ˆ

denote the prediction at time

*t*

*t*. Here we denote by *θ* the index

by T*n*(S*n*) is upper bounded in the following.

of an expert, taking value in an index set Λ = {1*,*2*,*··· *,N*} and

**Assumption 1.** For any target sequence T*n*, it holds that

*N* ∈R+ is the number of experts in the class. The performance

E [*L*(Tˆ (S )*,*T )] ≤ *C* (*n*)*,* (5)

of the predictions is evaluated by a non-negative loss function

*ℓ* :D×X →R+. We assume that only the forecaster has access to

where *CS*

*n n n S*

S*n*

(*n*) is a finite value depending on *n* for *n<*∞.

the SSI which may provide extra information on target sequence,

which is denoted by S*n* =(*XS,XS,...,XS*) for each *XS* ∈X. At

*t*

Assumption 1 does not significantly restrict the SSI as we only

ˆ

each time

1 2 *n*

*T* require that the total loss induced by T*n*(S*n*)∈X is bounded and

*t*, the forecaster predicts the current target instance *Xt*

we do not specify the function *C*

(*n*) at this stage. Clearly, *C*

(*n*)

with previous observations (*XT ,...,XT*

) up to time *t*−1 and the *S S*

1

corresponding SSI *S*

*t*

*t*−1

. In other words, the

depends the SSI through the conditional distribution *P* (*XS*|*XT* ),

prediction ˜

(*X*1 *,...,XS*) up to time *t*

*t t*

for which we will give two concrete examples in Section 3.

*Xt* can be regarded as a function of both SSI and target

With definitions in place, we introduce Algorithm 1, which

sequences *X*˜*t*(*XS,XS,...,XS,XT ,XT ,...,XT* ). We also use

1 2

˜ ˜ ˜

*t* 1 2

*t*−1

we call *Exp3 with SSI*. This algorithm is an extension of the

*n* :=(*X*1*,X*2*,...,X*˜*n*) to denote the sequence of the predictions.

T

Following the common assumption in the literature [8, 14] for deterministic online learning problems, we assume that the target sequence is an arbitrary sequence. It can even be viewed

classical Exponentially Weighted Average (Exp3) algorithm [8], which uses an exponentially updated mixture of the experts as the forecaster. Our algorithm further treats the maximum likelihood estimator *X*ˆ *T* (*XS*) as an additional expert, so that the prediction

as adversarially chosen by the “environment” with the knowledge *t t*

of the prediction rule of the forecaster. However, we assume the SSI is generated in a conditional independent stochastic fashion

made by the forecaster will partially depend on the information provided by the SSI.

by *Pn*(S*n*|T*n*)=Q *P* (*XS*|*XT* ) with some (known) conditional

*t*

*t*

*t*

**Algorithm 1** Exp3 with SSI

probability distribution *P* (*X*|*Y* ).

1: Initialize the weights for the SSI *wS* and all experts *wθ* to

To evaluate the performance of the prediction sequences, we firstly define the cumulative loss *L* which takes two sequences *An* :=(*a*1*,a*2*,...,an*) and *Bn* :=(*b*1*,b*2*,...,bn*) with length *n* as:

1 1

be 1;

2: **for** *t*=1 to *n* **do**

3: Receive *XS*;

*t*

4: *X* = *wSX*ˆ *T* (*XS*)+Σ*N*

˜

*t t t θ*=1 *t t*

*n*

*n*

*t*

*t*

*t*

5: Receive *XT* ;

*n*

*θ*=1

*wθfθ* ;

*t*

*L*(*A ,B* )=Σ*ℓ*(*a ,b* ) (1)

*t*=1

*t wS* +Σ*N wθ*

*t*

*θ*

*T*

6: The experts suffer the loss *ℓ*(*ft ,Xt* );

7: The SSI suffers the loss *ℓ*(*X*ˆ *T* (*XS*)*,XT* );

We use the absolute loss *ℓ*(*a,b*)=|*a*−*b*| throughout this paper in or-

*t t t*

ˆ *T*

*t*

der to derive the lower bounds [8]. We then define the *regret* for the

*S*

*t*+1

8: *w*

=*wSe*−*ηℓ*(T*n*(S*n*)*,Xt* );

deterministic online learning problems as the difference between

9: **for** *θ* =1 to *N* **do**

the cumulative loss between our forecaster and the best expert:

10: *wθ*

=*wθe*−*ηℓ*(*fθ,XT* )

*L*(T˜*n,*T*n*)−min*L*(F*θ,*T*n*)*.* (2)

*t*+1 *t t t*

11: **end for**

12: **end for**

*n*

*θ* 13: **return** T˜*n*

=(*X*˜

1*,X*˜

2*,...,X*˜*n*)

# Minimax Expected Regret

In this section, we consider a problem of minimising the expected

Since the decision space is convex and nonempty, the prediction *X*˜*t* in Algorithm 1 formed by a linear combination

D

regret for a worst-case target sequence , i.e. that maximises the expectation (w.r.t the side sequence) of (2). To this end, we will study the *minimax expected regret* defined as follows. Unless

*n*

*n*

T T

of the expert *fθ* and *X*ˆ *T* (*XS*) is also guaranteed to lie in the

decision space . In the following theorem, we give an upper bound on the minimax regret with the proposed algorithm.

D

*t*

*t*

*t*

**Theorem 1.** With Assumption 1, the minimax expected regret in (3) is upper bounded by:

The learning rate in its current form is not determined since the rate of *CS*(*n*) and *L*∗(*n*) may vary across different cases. In

Section 3, we will provide two specific structures for *C* (*n*) and

*R*(*n*)≤

2 log(*N* +1)+min{*CS*(*n*)−*L*∗(*n*)*,*0}*,* (6)

*L* (*n*) with two examples. Notably, Theorem 1 indicates that the effect of the SSI will be determined by the difference between

r*n* ∗ *S*

where we define *L*∗(*n*)=infmin*L*( *θ, n*).

*n*

F T

T*n θ*

*Proof.* The proof is different compared with the previous work [7, 8] that the prediction *pt* now depends on the target outcome *XT* by referencing the advice from both experts and the SSI *XS*.

*CS*(*n*) and *L*∗(*n*). In particular, if the expected cumulative loss of the side information is smaller than the loss induced by the best expert, the SSI is indeed helpful for predicting the target sequence. In contrast, when the SSI induces a higher loss compared to the best expert, the regret is upper bounded by *n*log(*N* +1),

*t t*

First of all, we lower bound the first term in (3) as follows.

infsup E {*L*(Ξ*n*(S*n,*T*n*)*,*T*n*)−min*Lθ*}

Ξ*n* T*n* S*n θ*

=infsup E {Σ|*pt*−*XT* |−minΣ|*fθ* −*XT* |} (7)

Ξ*n* T*n*

S*n*

*t*

*θ*

*t*

*t*

*t*

*t*

2

which is essentially the same as the learning bound without SSI

[7, 8, 9] when N is large. As a result, the second term in theorem 1 indicates how much the SSI can improve on the regret. To examine the tightness of the proposed upper bound, we also develop a lower bound for a particular outcome space and a

√

decision space in the following theorem.

1. Σ *T* Σ *θ T*

≥ {

*θ*

inf E E

Ξ*n* T*n*S*n*

|*pt*−*Xt* |−min

*t*

|*ft* −*Xt* |} (8)

*t*

**Theorem 2.** Consider X ={0*,*1} and D=[0*,*1], with the absolute

loss *ℓ*(*x,y*)=|*x*−*y*|, we have

=inf E E Σ|*pt*−*XT* |− EminΣ|*fθ* −*XT* |*,* (9) r

Ξ*n* T S

*t*

T

*θ*

*t*

*t*

1

*n*

*n*

*t*

*n*

*t*

∗

where inequality (a) holds since the worst-case target sequence will generate no lower regret than compared with any other

*R*(*n*)≥

log(*N* +1)+

2

*n*

*ξ* − 2

*n,* (15)

where *ξ*∗ =inf ˜ E*Z,XS* |*X*˜1(*XS*)−*Z*|, in which *Z* and *XS* are

*X*1

1

1

1

distribution, from which the target instances are i.i.d. Bernoulli

variables, i.e., *XT* ∼Ber(1 ). Under this case, the prediction made

stochastic target sequences. We now consider one potential

2

jointly distributed according to *P* (*Z,XS*), and *Z* is marginally

Bernoulli distributed as *Z* ∼Ber(1 ), *XS* is generated according

1

*t* 2 1

to the conditional distribution *S*

by the experts will always suffer an expected cumulative loss

of 1 *n*. Moreover, results from the independence between the target instances, and the side information is generated conditioned on those target instances, in which case the problem becomes *n* repetitive one-instance prediction problem:

2

inf E E

|*pt*−*X* |=*n*inf E

E |*p*1−*X* | (10)

*P* (*X*1 |*Z*).

*Proof.* (sketch) Under the absolute loss, we firstly lower bound the minimax regret *R*(*n*) as:

"Σ Σ #

*t*

*θ*

*t*

|*ft* −*Xt* |*,* (16)

T˜

*t*

*θ T*

Σ *R*(*n*)=infsup E

*T*

*T*

T

S

*t*

*t*

*t*

1

*n*

*n*

*n*

*t*

|*X*˜*t*(*XS*)−*XT* |−min

|*f* −*X* |

*n* T*n*S*n t*

Ξ

*p*

*T*

*S*

1 *X*1 *X*1

∗ *T*

E

|*Xt*(*Xt* )−*Xt* |− Emin

T˜*n* T*n,*S*n*

*X*1 *X*1

(*a*)

Σ ˜ *S T*

*t*

Σ *θ T*

T*n*

*θ*

*t*

It is known that for the absolute loss, the optimal forecaster *p*∗1 is determined by minimising:

=*n* E*T* E*S* |*p*1 −*X*1 |*,* (11)

≥ inf

Σ|*pt*−*X* |*P* (*X* |*X* )*.* (12)

*T T S*

1 1 1

*T*

*X*

1

where inequality (a) holds since the worst-case target sequence will generate no lower regret than compared with any other stochastic target sequences. We now assume that the target

instances and the SSI instances are generated according to a joint distribution *P* (*XS,XT* ) = *P* (*XT* )*P* (*XS*|*XT* ), here *P* (*XT* ) is

*t*

*t*

*t*

*t*

*t*

*t*

*t*

∼

Then, we denote by *ξ*∗ the expected loss induced by the *p*∗ in

1

eq (17). Note that, the optimality of *p*∗1 is w.r.t. the individual loss

|*p*∗1 −*XT* |. Following the equation (16), we have,

1

infsup E {*L*(*p,*T*n*)−min*Lθ*}

a Bernoulli distribution with probability 1 , i.e., *XT* Ber(1 ).

Clearly, in this case the expected loss incurred by the expert cannot be smaller than *n/*2. Then the sequential prediction problem becomes *n* repetitive one-instance prediction problem as follows:

2

2

Ξ*n* T*n* S*n θ*

!

E

≥

*nξ*∗− 1 *n*

+

1

*n*− Emin

|*f* −*X* |

Σ ˜ *S T*

T˜*n* T*n,*S*n*

*t*

˜∗ *S T*

*X*1 *,X*1

inf

Σ

*θ T*

(13)

≥

*ξ*∗−

*n*+

log*N,* (14)

*T* |*X*1(*XS*)−*XT* |*P* (*XT* |*XS*)

*X*1

1

1

1

1

*t*

*t*

|*Xt*(*Xt* )−*Xt* |=*n*

*T*E *S* |*X*1 (*X*1 )−*X*1 |*.* (17)

2 2 T*n θ t*

It is known that for the absolute loss, the optimal forecaster ˜∗

1 r*n* ˜ *X*1

Σ

where *P* (*XT* |*XS*) is induced by the joint distribution *P* (*XT ,XS*).

is determined by minimising

2

2

1

1

1

1

where the last step is derived with the same procedures from

Theorem 3.7 in [8].

Then we denote by *ξ*∗ the expected loss induced by *X*˜1∗ in (17),

and note that the optimality of *X*˜1∗ is w.r.t. the individual loss

|*X*˜1∗(*XS*)−*XT* |. Following (16), we have,

with the flipping probability *δ* with the target sequence being the

1 1

input. That is,

*R*(*n*)≥ *nξ*∗− 1 *n* + 1 *n*− EminΣ|*fθ* −*XT* |! (18)

*T S S*

2 2

≥

*ξ*∗− 2

*n*+

log*N,* (19)

2

1 r*n*

*t t*

T*n θ t*

*P* (*Xt* =*Xt* |*Xt* )=1−*δ,* (22)

*P* (*XT* =1−*XS*|*XS*)=*δ.* (23)

*t*

*t*

*t*

where the last step is derived with the similar procedures from Theorem 3.7 in [8].

It can be shown that the ML estimator Tˆ*n*(S*n*)=S*n* when *δ<* 1 ,

and Tˆ*n*(S*n*) = *S*¯*n* when *δ >* 1 , where *S*¯*n* := (1 − *XS,* 1 −

2

2

1

*XS,...,*1 − *XS*) denotes the sequence consisting of the flipped

T S

2

*n*

SSI instances. For *δ* = 1 , Tˆ*n*(S*n*) can be any value in D*n*. With

It can be easily checked that the first term in (15) is always 2 ˆ

negative since *ξ* is always smaller than 1 . So for large *n*, the lower bound is negative, showing that the loss produced by our forecaster could potentially be much smaller than the best expert. It can also be seen that if the term *Cs*(*n*) *L*∗(*n*) in the upper bound takes the form *cn* for some positive *c*, then the upper and lower bound are matched in terms of the scaling law. In the next section, we show two examples demonstrating this point.

2

−

−

# EXAMPLES

the maximum likelihood estimator *n*( *n*), we can calculate the expected loss as:

E [*L*(Tˆ*n*(S*n*)*,*T*n*)] =*n* *δ*∧¯*δ* (24)

S*n*

where *a*∧*b* =min{*a,b*} and ¯*δ* =1−*δ*, which satisfies Assumption 1 with *CS*(*n*)=*n δ*∧¯*δ* .

**Corollary 1.** Under the binary flipping channel setup, when

*L*∗(*n*) grows linearly in *n*, i.e. *L*∗(*n*)=*cfn*, we have

In this section, we consider two concrete online learning prob- r*n* ¯ }

lems and derive their corresponding upper and lower bounds to verify the effectiveness of the proposed bounds. To characterize

*R*(*n*)≤

log(*N* +1)+min

2

0*,n*(*δ*∧*δ*−*cf* )

*.* (25)

the behavior of the expert class, we will further consider the expert class generating a cumulative loss that scales linearly in *n*. The following expert class with an example displays one of the possible case satisfying the linear loss expert.

**Definition 1.** (Constant Expert) We say an expert class is the

*Proof.* The proof is straightforwardly following Theorem 1 from equation (**??**)

r*n*

*R*(*n*)=

log(*N* +1)

2

constant expert class such that all experts in the class yield a fixed

+sup{min{ E (*L*(Tˆ*n*(S*n*)*,*T*n*)−min*Lθ*)*,*0}} (26)

prediction for any target instances. Mathematically,

For all *t* from 1 to *n*, *fθ* =*cθ,* (20)

*t*

where

T*n* S*n*

r*n*

=

2 log(*N* +1)+min{0*,n δ*∧*δ*−*cf*

¯

*cθ* is some constant in D.

*θ*

(27)

}*.* (28)

Since the constant expert class is independent to the target instances, we can directly calculate the amount *L*∗(*n*) defined in Theorem 1 under a certain setup for the decision space D and

Notice that if any expert in the expert class F*θ* =(*fθ,fθ,...,fθ*)

*n*

1

2

*n*

output space X. We give two examples as follows.

**Example 1.** Assume the decision space is D = [0*,*1], and we

suffers a cumulative loss more than 1 *n*, one can construct a new

expert class (1 *fθ,*1 *fθ,...,*1 *fθ*) that suffers a loss smaller

2

1

2

*n*

— − −

than 1 *n*. Hence we only consider the case that *cf* is always smaller

consider a constant expert class that each expert predict a fixed 2 1 ¯

*c*

D

∧

constant *fθ* in ]. Also, assume that there always exists two experts predicting 0*.*1 and 0*.*7 for any time *t*. We also consider the binary output space, e.g, X ={0*,*1}. Then we have

or equal to 2 . From Corollary 1, we notice that if *δ δ* is smaller

than *cf* , the regret is asymptotically negative and scale linearly with *n*. In the following, we give the corresponding lower bound.

**Corollary 2.** Under the binary symmetric channel setup, we have

*L*∗(*n*)=infmin*L*(F*θ,*T*n*)=0*.*1*n.* (21)

*R*(*n*)≥

2 log(*N* +1)−*n*

2 −*δ*∧¯*δ*

*.* (29)

T*n θ*

∧

**3.1. SSI via a binary symmetric channel**

r*n*  1

In this example, we consider a learning problem setup that

= 0*,*1 *,* =[0*,*1] under the absolute loss *ℓ*(*x,y*)= *x y* . We assume that the SSI is the output of a binary symmetric channel

X { } D | − |

We see that in the case when *δ* ¯*δ < cf* , the upper and the lower bound is matched in terms of the scaling law of order Ω(−*n*) (although with a different constant).

# 3.2. SSI via a Zero-mean Gaussian Channel

Now we consider a different type of side information such that the side instance is the noisy version of the target instance pair-wise:

It can be seen that the upper bound in this example behaves similarly to that in the binary symmetric channel case. When the quantity Φ(− 1 ) is smaller than *cf* , the upper bound of the

*X* = *X* +*N* , where *N* ∼ N(0*,σ*2). Here we still assume

minimax regret becomes negative with a large *n*. Intuitively,

*S,t*

*T,t t t*

when *σ* is large, the quantity Φ(− 1 ) will become larger, which

the target instances are restricted in a binary outcome space 0*,*1 . Note that in this problem setup, the side instances are drawn from a distribution over the space *R*, which differs from the instance space X.

2*σ*

{ }

2*σ*

decreases the effectiveness of the SSI.

**Corollary 4.** Under the zero-mean Gaussian channel setup, we have

We can easily determine the maximum likelihood estima-

2

2*σ*

2

tor *X*ˆ *T* (*XS*) for this problem: *X*ˆ *T* (*XS*) = 1 when *XS* ≥ 1 ,

*t*

*t*

*t*

*t*

*t*

2

and *X*ˆ *T* (*XS*) = 0 when *XS*

*<* 1 . By introducing the cu-

*R*(*n*)≥r*n*log(*N* +1)+*n* Φ(− 1 )− 1 *.* (37)

mulative density function of the standard normal distribution

*t*

*t*

*t*

2

√1 ∫ *z*

2

Φ(*z*)=

−∞*e*

*dt*, we have

2*σ*

bound will become negative when *n*

2

2*π*

−*t /*2

Similarly, as *σ>* 0, we have Φ(− 1 ) *<* Φ(0) = 1 , the lower

1

E [*L*(Tˆ (S )*,*T )] =*n*Φ(− )*.* (30)

*n n n*

2*σ*

S*n*

**Corollary 3.** Under the zero-mean Gaussian channel setup, and when *L*∗(*n*) is linear to *n*, i.e. *L*∗(*n*)=*cfn*, we have

*R*(*n*)≤r*n*log(*N* +1)+min{0*,n* Φ(− 1 )−*c*  } (31)

2

2*σ*

*f*

*Proof.* Following the proof of Theorem 2, we need to specify the quantity *ξ*∗. We start by finding the optimal forecaster *p* for pre-

increases. Similar to the

binary symmetric channel example, the upper and the lower bound is matched in terms of the scaling law if *cf >*Φ(− 1 ).

2*σ*

# CONCLUSION AND FUTURE WORKS

This work shows the upper and lower regret bounds on general deterministic online learning problems with two concrete exam-

ples, where an additional stochastic sequential side information

sequence is revealed to the forecaster. The result infers the effectiveness of the the side information which may significantly

dicting the target instances *T*

*t* improved the learning rate and shows the possibility of producing

*Xt* by minimising the absolute loss:

E *l*(*p* (*XS*)*,XT* )=ΣΣ|*p* (*XS*)−*XT* |*P* (*XS,XT* )

*X*

*T*

*t*

*t*

*t*

*t*

*t*

*t*

*t*

*t*

*t*

*t*

*t*

*XT XS*

a negative regret. For future works, one may wish to find a tighter lower bound on the minimax regret based on more advanced

algorithms, or more elementary proofs.

=Σ Σ|*pt*(*XS*)−*XT* |*P* (*XT* |*XS*)

*S*

*t*

*T*

*t*

*t*

*t*

*t*

*t*

*X*

*X*

*P* (*XS*)*.* (32)

**References**

Then minimising the expected loss w.r.t. *pt*(*XS*) is equivalently

*t*

[1] N. Merhav and M. Feder, “Universal prediction,” *IEEE*

*Transactions on Information Theory*, vol. 44, no. 6, pp.

minimising Σ

*t*

*t*

*t*

*S T T S*

*S*. Given

2124–2147, 1998.

*XT* |*pt*(*Xt* )−*Xt* |*P* (*Xt* |*Xt* ) for any *Xt*

*XS* =1, we have

Σ|*p* (*XS* =1)−*XT* |*P* (*XT* |*XS* =1) (33)

*X*

*T*

*t*

*t*

*t*

*t*

*t*

*t*

[2] X. Wu, J. H. Manton, U. Aickelin, and J. Zhu, “Online

transfer learning: Negative transfer and effect of prior

knowledge,” in *2021 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2021, pp. 1540–1545.

=(1−*δ*)|*pt*(*XS* =1)−1|+*δ*|*pt*(*XS* =1)−0|*.* (34)

*t t* [3] V. Vovk, “A game of prediction with expert advice,” *Journal*

Then we can obtain when *δ <* 1 , *pt*∗(*XS*) = *XS*, when *δ >* 1 ,

*of Computer and System Sciences*, vol. 56, no. 2, pp.

2 *t t* 2

*p*∗*t* (*XS*) = *XS*, and when *δ* = 1 , there are an infinite number of

153–173, 1998.

*t t* 2

minimizers *pt*∗(*XS*) between 0 and 1. One can verify that the optimal forecaster *pt*∗(*XS*) is the maximum likelihood estimator Tˆ*n*(S*n*). Then similar to (24), we have

*t*

*t*

1. N. Littlestone and M. K. Warmuth, “The weighted majority algorithm,” *Information and computation*, vol. 108, no. 2, pp. 212–261, 1994.

1

*ξ*∗ = E E |*p*∗−*XT* |= E [*L*(Tˆ (S )*,*T )] (35)

*t*

1. V. G. Vovk, “Aggregating strategies,” *Proc. of Computational*

*t*

*XT XS*

*t*

*t*

*n*S*n*

*n n n*

*Learning Theory, 1990*, 1990.

=*δ*∧¯*δ.* (36)

By substituting the *ξ*∗ in Theorem 2 as (36), we completed the proof.

1. P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, “The nonstochastic multiarmed bandit problem,” *SIAM journal on computing*, vol. 32, no. 1, pp. 48–77, 2002.
2. D. Haussler, J. Kivinen, and M. K. Warmuth, “Tight worst-case loss bounds for predicting with expert advice,” in *European Conference on Computational Learning Theory*. Springer, 1995, pp. 69–83.
3. N. Cesa-Bianchi and G. Lugosi, *Prediction, learning, and games*. Cambridge university press, 2006.
4. N. D. Vanli and S. S. Kozat, “A unified approach to universal prediction: Generalized upper and lower bounds,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 3, pp. 646–651, 2014.
5. T. M. Cover and E. Ordentlich, “Universal portfolios with side information,” *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 348–363, 1996.
6. Q. Xie and A. R. Barron, “Asymptotic minimax regret for data compression, gambling, and prediction,” *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 431–445, 2000.
7. A. Bhatt and Y.-H. Kim, “Sequential prediction under log-loss with side information,” in *Algorithmic Learning Theory*. PMLR, 2021, pp. 340–344.
8. L. Torrey and J. Shavlik, “Transfer learning,” in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*. IGI global, 2010, pp. 242–264.
9. N. Cesa-Bianchi and F. Orabona, “Online learning algo- rithms,” *Annual review of statistics and its application*, 2021.