Solution for Exercise 1:

Pick η_t as $\eta_t = \frac{\gamma}{M-m}\sqrt{\frac{\ln N}{t}}$ where γ is to be determined by the analysis

Hoeffding's lemma: $\delta_t \leq \frac{\eta_t}{8}(M-m)^2$

So that:
$$R_T = \sum_{t,j} p_{jt}\, \ell_{jt} - \min_k \sum_{t=1}^{T} \ell_{kt}$$
$$\leq \frac{\ln N}{\eta_T} + \sum_{t=1}^{T} \delta_t \leq \frac{M-m}{\gamma}\sqrt{T \ln N} + \gamma \frac{(M-m)}{8} \sum_{t=1}^{T} \sqrt{\frac{\ln N}{t}}$$

Now,
$$\sum_{t=1}^{T} 1/\sqrt{t} \leq \int_0^T \frac{1}{\sqrt{t}}\, dt = \left[2\sqrt{t}\right]_0^T = 2\sqrt{T}$$

$$R_T \leq (M-m)\sqrt{T \ln N}\, \underbrace{\left(\frac{1}{\gamma} + \frac{2\gamma}{8}\right)}_{\text{optimal value } \gamma^* = 2} = (M-m)\sqrt{T \ln N}$$

better constant than with the doubling trick

We only loose a $\sqrt{2}$ factor wrt bound when T is know
($\sqrt{2}$ is the price for adaptivity in T).

Solution for Exercise 2:

If you find one (even in several weeks!) please send it to me! The reward for a significant contribution will be given by bonus points at the exam.

Solution for Exercise 3.

(1)
$$v_t = \sum_j p_{jt}\Big(\ell_{jt} - \sum_k p_{kt}\ell_{kt}\Big)^2$$
is some variance thus, by the variational formula for the variance

$$= \min_{\mu\in\mathbb{R}} \sum_j p_{jt}(\ell_{jt}-\mu)^2$$

pick $\mu = 0$

$$\leqslant \sum_j p_{jt}\,\ell_{jt}^2$$

$\ell_{jt} \geqslant 0$ and $\ell_{jt} \leqslant M$ thus $\ell_{jt}^2 \leqslant M\ell_{jt}$

$$\leqslant M \sum_j p_{jt}\,\ell_{jt}$$

Hence, summing over $t=1,\dots T$:

$$\sum_{t=1}^T v_t = \sum_{t=1}^T \sum_{j=1}^N p_{jt}\Big(\ell_{jt} - \sum_{k=1}^N p_{kt}\ell_{kt}\Big)^2 \leqslant M \sum_{t=1}^T \sum_j p_{jt}\,\ell_{jt}$$

(2) The theorem for EWA tuned with $\eta_t = \dfrac{\ln N}{\sum_{s\leqslant t-1} \delta_s}$ ensures that for this algorithm (since $m=0$) :

$$\sum_{t,j} p_{jt}\,\ell_{jt} - \min_k \sum_t \ell_{kt} \leqslant \underbrace{2\sqrt{\sum_t v_t \ln N}}_{\leqslant 2\sqrt{M\sum_{t,j} p_{jt}\ell_{jt}\ln N}} + M\Big(2 + \frac{4}{3}\ln N\Big) \qquad (*)$$

Thus, denoting $x = \sqrt{\sum_{t,j} p_{jt}\ell_{jt}}$, we have the 2nd order inequality:

$$x^2 \leqslant \Big(2\sqrt{M\ln N}\Big)x + \Big(\min_k \sum_{t=1}^T \ell_{kt} + M\Big(2+\frac{4}{3}\ln N\Big)\Big)$$

A lemma that we saw in class says that if $x^2 \leqslant b + a\sqrt{x}$ $(a,b\geqslant 0)$ then $x \leqslant a + \sqrt{b}$

This means here that
$$\sqrt{\sum_{t,j} p_{jt}\ell_{jt}} \leqslant 2\sqrt{M\ln N} + \underbrace{\sqrt{\min_k \sum_t \ell_{kt} + M\Big(2+\frac{4}{3}\ln N\Big)}}_{\leqslant \sqrt{\min_k \sum_t \ell_{kt}} + \sqrt{M\big(2+\frac{4}{3}\ln N\big)}}$$

We could take $(\ )^2$ of both sides but it is slightly more elegant to substitute this inequality into $(*)$

We get:

$$\sum_{t,j} p_{jt}\, \ell_{jt} - \min_k \sum_t \ell_{kt}$$

$$\leq \underbrace{4M\ln N + 2M\sqrt{\left(2+\tfrac{4}{3}\ln N\right)\ln N} + M\left(2+\tfrac{4}{3}\ln N\right)}_{\text{let's make these more readable!}} + 2\sqrt{M \min_k \sum_t \ell_{kt} \ln N}$$

(3)

We may assume that $N \geq 2$, in which case $\ln N \geq 0.693$ and $1 \leq \frac{3}{2}\ln N$

then,

$$4M\ln N + 2M\sqrt{\underbrace{\left(2+\tfrac{4}{3}\ln N\right)}_{\leq 2\times\frac{3}{2}\ln N}\ln N} + M\underbrace{\left(2+\tfrac{4}{3}\ln N\right)}_{\leq 2\times\frac{3}{2}\ln N} \leq M\ln N \times \underbrace{\left(4 + 2\sqrt{3+\tfrac{4}{3}} + 3 + \tfrac{4}{3}\right)}_{\leq 13}$$

final bound:

$$\sum_{t,j} p_{jt}\, \ell_{jt} - \min_k \sum_{t=1}^{T} \ell_{kt} \leq 13M\ln N + 2\sqrt{M \min_k \sum_t \ell_{kt} \ln N}$$

error suffered when $\exists k \mid \sum_t \ell_{kt} = 0$ ie, all $\ell_{kt} = 0$ given that losses are non-negative

Solution for Exercise 4.

Assume that $\eta \leq 1/2M$: $-\eta \ell_{ks} \geq -1/2$, positive weights p_{jt}, algorithm well defined

(1) $$\underbrace{-\eta \sum_{j=1}^{N} p_{jt} \ell_{jt}}_{\geq -1/2} \underset{\ln(1+u) \leq u}{\geq} \ln\Big(1 - \eta \sum_j p_{jt} \ell_{jt}\Big)$$

$$= \ln\Big(\sum_j p_{jt}(1 - \eta \ell_{jt})\Big)$$

$$\underset{\text{def. of } p_{jt}}{=} \ln \frac{\sum_j \prod_{s=1}^{t} (1 - \eta \ell_{js})}{\sum_{k \leq N} \prod_{s=1}^{t-1} (1 - \eta \ell_{ks})}$$

Telescoping sum:

$$-\eta \sum_{t=1}^{T} \sum_{j=1}^{N} p_{jt} \ell_{jt} \underset{\substack{\text{convention:} \\ \text{empty } \prod \text{ equals } 1}}{\geq} \ln \frac{\sum_{k \leq N} \prod_{t=1}^{T} (1 - \eta \ell_{kt})}{N}$$

We lower bound $\ln\Big(\sum_{k \leq N} \prod_{t=1}^{T} (1 - \eta \ell_{kt})\Big)$:

$\forall j,$ $$\geq \ln\Big(\prod_{t=1}^{T} (1 - \eta \ell_{jt})\Big) = \sum_{t=1}^{T} \ln(1 - \eta \ell_{jt})$$

$$\geq -\eta \ell_{jt} - \eta^2 \ell_{jt}^2$$

$-\eta \ell_{jt} \geq -1/2$ and $\ln(1+u) \geq u - u^2$ $\forall u \geq -1/2$

Summarizing what we got so far:

$\forall j,$ $$-\eta \sum_{t=1}^{T} \sum_{k} p_{kt} \ell_{kt} \geq -\ln N - \eta \sum_{t=1}^{T} \ell_{jt} - \eta^2 \sum_{t=1}^{T} \ell_{jt}^2$$

Thus, $\forall j,$ $$\sum_{t=1}^{T} \sum_{k=1}^{N} p_{kt} \ell_{kt} - \sum_{t=1}^{T} \ell_{jt} \leq \frac{\ln N}{\eta} + \eta \sum_{t=1}^{T} \ell_{jt}^2.$$

We did so under the assumption that $\forall j, t,$ $-\eta \ell_{jt} \geq -1/2$ valid as soon as $\ell_{jt} \leq M$ and $\eta \leq 1/2M$.

(2) Waiting for your solutions! (But keep in mind that it's a difficult problem...)