Show that $E(\boldsymbol{a}) = \sum_{i \in I} log(1 + e^{-y_i \boldsymbol{a}^T \boldsymbol{x}_i})$ is a convex function of $\boldsymbol{a}$.

Proof.

$$
\begin{aligned}
\frac{\partial E}{\partial \boldsymbol{a}} &= \sum_{i \in I} \frac{\partial log(1 + e^{-y_i \boldsymbol{a}^T \boldsymbol{x}_i})}{\partial \boldsymbol{a}} \\
&= \sum_{i \in I} \frac{e^{-y_i \boldsymbol{a}^T \boldsymbol{x}_i}}{1 + e^{-y_i \boldsymbol{a}^T \boldsymbol{x}_i}} \cdot (-y_i) \cdot \boldsymbol{x}_i^T \\
&\triangleq - \sum_{i \in I} y_i \sigma(-y_i \boldsymbol{a}^T \boldsymbol{x}_i) \boldsymbol{x}_i^T
\end{aligned}
$$

Here the function $\sigma(\cdot)$ is sigmoid function, and $\frac{d\sigma(x)}{dx} = \sigma(x)(1-\sigma(x))$. Then we have

$$
\begin{aligned}
\frac{\partial^2 E}{\partial a_p a_q} &= \sum_{i \in I} y_i^2 x_{ip} x_{iq} \sigma(-y_i \boldsymbol{a}^T \boldsymbol{x}_i)(1 - \sigma(-y_i \boldsymbol{a}^T \boldsymbol{x}_i)) \\
&= \sum_{i \in I} x_{ip} x_{iq} \sigma(-y_i \boldsymbol{a}^T \boldsymbol{x}_i)(1 - \sigma(-y_i \boldsymbol{a}^T \boldsymbol{x}_i)) \\
&\triangleq \sum_{i \in I} x_{ip} x_{iq} \rho_i
\end{aligned}
$$

Then we need proof that the Hessian matrix $\nabla_{\boldsymbol{a}}^2 E$ is semidefinite, which equals $\forall \boldsymbol{v}, \boldsymbol{v}^T (\nabla_{\boldsymbol{a}}^2 E) \boldsymbol{v} \geq 0$.

$$
\begin{aligned}
\boldsymbol{v}^T (\nabla_{\boldsymbol{a}}^2 E) \boldsymbol{v} &= \sum_p \sum_q \sum_i v_p v_q x_{ip} x_{iq} \rho_i \\
&= \sum_i \sum_p \sum_q v_p v_q x_{ip} x_{iq} \rho_i \\
&= \sum_i \rho_i \sum_p v_p x_{ip} \sum_q v_q x_{iq} \\
&= \sum_i \rho_i \sum_p v_p x_{ip} \boldsymbol{x}_i^T \boldsymbol{v} \\
&= \sum_i \rho_i (\boldsymbol{x}_i^T \boldsymbol{v})^2 \geq 0
\end{aligned}
$$

Therefore, Hessian matrix $\nabla_{\boldsymbol{a}}^2 E$ is semidefinite, which indicates $E(\boldsymbol{a})$ is a convex function of $\boldsymbol{a}$.

$\square$