Several tasks in machine learning are evaluated using non-differentiable metrics such as mean average precision or Spearman correlation. However, their non-