Summary

Online Learning is a fundamental machine learning setting in which a learner is to sequentially issue predictions given some (partial) knowledge about previous correct predictions and possibly additional information. The environment of the learner is often assumed to be adversarial, making the learner's task of suffering as little loss as possible difficult. Nevertheless, in the last three decades many different Online Learning algorithms have been successfully shown to provide satisfying guarantees in various settings. The guarantees in Online Learning are about regret, which is the difference between the cumulative loss of the learner and the cumulative loss the offline optimizer of the loss, which is also known as the comparator. In this dissertation we provide several new insights in many different settings of Online Learning, hence the title of the dissertation.

In Chapter 2 we study one of the most fundamental algorithms in Online Learning: Exponential Weights. We show how to tune Exponential Weights such that it can be applied to several different settings and show that with specific parameter choices we recover several other important algorithms in Online Learning as special cases of Exponential Weights. This provides a centralized understanding of many algorithms in the Online Learning setting and unifies the analysis of these algorithms.

An important distinction in Online Learning is between the full-information and bandit settings. In the full-information setting the environment reveals all information to the learner but in the more challenging bandit setting the environment only reveals partial information. An important property of many algorithms in both the full-information and bandit settings is that they are able to provide suitable regret bounds even in adversarial environments. However, these algorithms are often tuned to only deal with adversarial environments and are not able to exploit more benign environments. A recurring subject in this dissertation is how to design algorithms that are able to exploit benign environments but also provide suitable guarantees in adversarial environments, without knowing what type of environment the learner faces beforehand. These algorithms are known as adaptive algorithms as they adapt to the environment. In Chapter 2 we show how we can recover several adaptive algorithms as special cases of Exponential Weights. In Chapters 3 and 4 we study a particular type of adaptive algorithms, namely comparator-adaptive algorithms. The regret bounds of comparator-adaptive algorithms depend on properties of the offline

minimizer of the loss, which in some cases can lead to smaller regret compared to the regret of standard algorithms. In Chapter 3 we show how we can modify comparator-adaptive algorithms to adapt to unknown noise, which is useful when people want to choose how much privacy they have without disclosing how much they value their privacy. Additionally, when the losses are nice in a particular sense we show that our modified comparator-adaptive algorithm has low regret. In Chapter 4 we provide the first comparator-adaptive algorithms for the Bandit Convex Optimization setting. This can be especially advantageous when the comparator is small when measured in a particular norm, as this leads to smaller regret bounds compared to non-adaptive algorithms. In Chapter 5 we present MetaGrad, which adapts to a broad class of functions. The class of functions to which MetaGrad is adaptive includes exp-concave losses, losses with unknown lipschitz constants, and various other types of stochastic or non-stochastic functions.

Since the learner updates his prediction in each round an important property of Online Learning algorithms is the running time. The per round running time is often considered too high to be practical whenever the updates take more than quadratic time in the dimension of the problem. Because of this reason considerable effort has been made to improve the running time of many Online Learning algorithms, including in this dissertation. In Chapter 6 provide a new algorithm with often similar or better guarantees than slower algorithms for the Online Multiclass Classification setting. In this setting in each round the learner has to predict a label given a d-dimensional feature vector which contains additional information. In the Online Multiclass Classification setting the learner suffers the zero-one loss, which is one whenever the learner makes a mistake and zero whenever he correctly predicts the label. The benchmark in the Online Multiclass Classification setting is a convex surrogate loss which upper bounds the zero-one loss and the goal of the learner is to minimize the surrogate regret: the difference between the sum of the zero-one losses and the offline minimum of the sum of the surrogate losses. Previous algorithms in the Online Multiclass Classification setting often relied on second-order algorithms to guarantee small regret. Second-order algorithms keep track of a d by d matrix of parameters, which the learner updates in each round, making the per round running time at least d^2. We introduce a novel algorithm called GAPTRON which has a per round running time of order $O(d)$. Surprisingly, GAPTRON often matches or improves upon the guarantees of slower algorithms. For example, in the Bandit Multiclass Classification setting the surrogate regret bound of GAPTRON is a factor $\sqrt{d}$ smaller than slower algorithms. We achieve our results by using a new approach which exploits the gap between the zero-one loss and a surrogate loss. This new approach allows the learner to use a linear time algorithm to update the predictions while still obtaining small surrogate regret

bounds. Other improvements in running time of algorithms are made in Chapter 5, in which we show how to improve the running time of the aforementioned adaptive algorithm MetaGrad by using sketching methods. In Chapter 7 we consider online portfolio selection. The optimal algorithm for online portfolio selection is a version of Exponential Weights. Unfortunately the running time of this particular version of Exponential Weights is too high to be considered practical. Many different algorithms for online portfolio selection have been considered, all of them with different shortcomings. We pose an open problem which asks for a fast and optimal algorithm and provide the first steps of the analysis of an algorithm we think is the answer to the open problem. We then show that in particular cases the proposed algorithm indeed yields the optimal regret bound.