

# Table of Content

<b>1.</b>	<b>What is Machine Learning and why it is important.</b>	<b>1</b>
<b>2.</b>	<b>Bayesian Autoregressive Spectral Estimation</b>	<b>3</b>
2.1.	Introduction . . . . .	3
2.2.	Background: Parametric spectral estimation using an AR model . . . . .	4
2.2.1.	Power spectral density of AR( $p$ ) processes . . . . .	4
2.3.	Bayesian spectral estimation for the autoregressive model . . . . .	5
2.3.1.	Generic Bayesian parametric spectral estimation . . . . .	5
2.3.2.	Establishing a prior over AR models . . . . .	6
2.3.3.	Model likelihood . . . . .	7
2.3.4.	Closed form posterior . . . . .	8
2.4.	Simulations . . . . .	9
2.4.1.	Misspecified model order: AR-generated time series . . . . .	9
2.4.2.	Continuous-time signal: Gaussian process with Laplace covariance . . . . .	10
2.4.3.	Finding the periodicity of an astronomical time series . . . . .	11
2.4.4.	Performance of the closed-form posterior . . . . .	13
2.5.	Conclusion . . . . .	15
2.6.	Future work . . . . .	16
<b>3.</b>	<b>Neural networks trained with Cross-entropy loss follow the NTK</b>	<b>17</b>
3.1.	Introduction . . . . .	17
3.2.	Background . . . . .	18
3.2.1.	Neural Networks . . . . .	19
3.2.2.	Kernel Gradient . . . . .	19
3.2.3.	Neural tangent kernel . . . . .	20
3.3.	Cross entropy loss function . . . . .	21
3.3.1.	$C$ is convex and non negative . . . . .	22
3.3.2.	Direction of training $d_t$ . . . . .	22
3.3.3.	$\int_0^T \ d_t\ _{p_{in}} dt$ stays stochastically bounded. . . . .	23
3.4.	Experiments . . . . .	24
3.4.1.	Convergence of the empirical NTK of the neural network to a fixed limiting kernel . . . . .	25
3.4.2.	Convergence of the finite neural network to the infinite-width limit . . . . .	29
3.5.	Conclusion . . . . .	30
3.6.	Future work . . . . .	33
<b>Bibliography</b>		<b>35</b>